



Forward-Backward Stochastic Differential Equations and Controlled McKean Vlasov Dynamics

René Carmona, François Delarue

► To cite this version:

René Carmona, François Delarue. Forward-Backward Stochastic Differential Equations and Controlled McKean Vlasov Dynamics. *Annals of Probability*, 2015, 43, pp.2647-2700. hal-00803683v2

HAL Id: hal-00803683

<https://hal.science/hal-00803683v2>

Submitted on 9 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND CONTROLLED MCKEAN VLASOV DYNAMICS

RENÉ CARMONA AND FRANÇOIS DELARUE

ABSTRACT. The purpose of this paper is to provide a detailed probabilistic analysis of the optimal control of nonlinear stochastic dynamical systems of the McKean Vlasov type. Motivated by the recent interest in mean field games, we highlight the connection and the differences between the two sets of problems. We prove a new version of the stochastic maximum principle and give sufficient conditions for existence of an optimal control. We also provide examples for which our sufficient conditions for existence of an optimal solution are satisfied. Finally we show that our solution to the control problem provides approximate equilibria for large stochastic games with mean field interactions.

1. INTRODUCTION

The purpose of this paper is to provide a detailed probabilistic analysis of the optimal control of nonlinear stochastic dynamical systems of the McKean-Vlasov type. The present study is motivated in part by a recent surge of interest in mean field games.

We prove a version of the stochastic Pontryagin maximum principle that is tailor-made to McKean-Vlasov dynamics and give sufficient conditions for existence of an optimal control. We also provide a class of examples for which our sufficient conditions for existence of an optimal solution are satisfied. Putting these conditions to work at the solution of an optimal control problem leads to the solution of a system of Forward Backward Stochastic Differential Equations (FBSDEs for short) where the marginal distributions of the solutions appear in the coefficients of the equations. We call these equations mean field FBSDEs, or FBSDEs of McKean-Vlasov type. To the best of our knowledge, these equations have not been studied before. A rather general existence result was recently proposed in [6], but one of the assumptions (boundedness of the coefficients with respect to the state variable) precludes applications to solvable models such that the Linear Quadratic (LQ for short) models. Here, we take advantage of the convexity of the underlying Hamiltonian and apply the so-called *continuation method* exposed in [14] in order to prove existence and uniqueness of the solution of the FBSDEs at hand, extending and refining the results of [6] to the models considered in this paper. The technical details are given in Section 5.

Without the control α_t (see Eq. (1) right below), stochastic differential equations of McKean-Vlasov type are associated to special nonlinear Partial Differential Equations (PDEs) put on a rigorous mathematical footing by Henry McKean Jr in [11]. See also [12, 16, 10]. Existence and uniqueness results for these equations have been developed in order to provide effective equations for studying large systems, reducing the dimension and the complexity, at the cost of handling non Markovian

Date: June 10, 2011.

2000 Mathematics Subject Classification. Primary .

Partially supported by NSF: DMS-0806591.

dynamics depending upon the statistical distribution of the solution. In the same spirit, we show in Section 6 that our solution of the optimal control of McKean-Vlasov stochastic differential equations provides strategies putting a large system of individual optimizers in an approximate equilibrium, the notion of equilibrium being defined appropriately. The proof is based on standard arguments from the theory of the propagation of chaos, see for example [16, 10]. The identification of approximate equilibria in *feedback form* requires strong regularity properties of the decoupling field of the FBSDE. They are proved in Section 5.

2. PROBABILISTIC SET-UP OF MCKEAN-VLASOV EQUATIONS

In what follows, we assume that $W = (W_t)_{0 \leq t \leq T}$ is an m -dimensional standard Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ is its natural filtration possibly augmented with an independent σ -algebra \mathcal{F}_0 . For each random variable/vector or stochastic process X , we denote by \mathbb{P}_X the law (alternatively called the distribution) of X .

The stochastic dynamics of interest in this paper are given by a stochastic process $X = (X_t)_{0 \leq t \leq T}$ satisfying a nonlinear stochastic differential equation of the form

$$(1) \quad dX_t = b(t, X_t, \mathbb{P}_{X_t}, \alpha_t)dt + \sigma(t, X_t, \mathbb{P}_{X_t}, \alpha_t)dW_t, \quad 0 \leq t \leq T,$$

where the drift and diffusion coefficient of the state X_t of the system are given by the pair of deterministic functions $(b, \sigma) : [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^d \times \mathbb{R}^{d \times m}$ and $\alpha = (\alpha_t)_{0 \leq t \leq T}$ is a progressively measurable process with values in a measurable space (A, \mathcal{A}) . Typically, A will be an open subset of an Euclidean space \mathbb{R}^k and \mathcal{A} the σ -field induced by the Borel σ -field of this Euclidean space.

Also, for each measurable space (E, \mathcal{E}) , we use the notation $\mathcal{P}(E)$ for the space of probability measures on (E, \mathcal{E}) , assuming that the σ -field \mathcal{E} on which the measures are defined is understood. When E is a metric or a normed space (most often \mathbb{R}^d), we denote by $\mathcal{P}_p(E)$ the subspace of $\mathcal{P}(E)$ of the probability measures of order p , namely those probability measures which integrate the p -th power of the distance to a fixed point (whose choice is irrelevant in the definition of $\mathcal{P}_p(E)$). The term *nonlinear* does not refer to the fact that the coefficients b and σ could be nonlinear functions of x but instead to the fact that they depend not only on the value of the unknown process X_t at time t , but also on its marginal distribution \mathbb{P}_{X_t} . We shall assume that the drift coefficient b and the volatility σ satisfy the following assumptions.

(A1) For each $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\alpha \in A$, the function $[0, T] \ni t \mapsto (b, \sigma)(t, x, \mu, \alpha) \in \mathbb{R}^d \times \mathbb{R}^{d \times m}$ is square integrable;

(A2) $\exists c > 0, \forall t \in [0, T], \forall \alpha \in A, \forall x, x' \in \mathbb{R}^d, \forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d),$

$$|b(t, x, \mu, \alpha) - b(t, x', \mu', \alpha)| + |\sigma(t, x, \mu, \alpha) - \sigma(t, x', \mu', \alpha)| \leq c[|x - x'| + W_2(\mu, \mu')];$$

where $W_2(\mu, \mu')$ denotes the 2-Wasserstein distance. For a general $p > 1$, the p -Wasserstein distance $W_p(\mu, \mu')$ is defined on $\mathcal{P}_p(E)$ by:

$$W_p(\mu, \mu') = \inf \left\{ \left[\int_{E \times E} |x - y|^p \pi(dx, dy) \right]^{1/p}; \pi \in \mathcal{P}_2(E \times E) \text{ with marginals } \mu \text{ and } \mu' \right\}.$$

Notice that if X and X' are random variables of order 2, then $W_2(\mathbb{P}_X, \mathbb{P}_{X'}) \leq [\mathbb{E}|X - X'|^2]^{1/2}$.

The set \mathbb{A} of so-called *admissible* control processes α is defined as the set of A -valued progressively measurable processes $\alpha \in \mathbb{H}^{2,k}$, where $\mathbb{H}^{2,n}$ denotes the Hilbert space

$$\mathbb{H}^{2,n} := \left\{ Z \in \mathbb{H}^{0,n}; \mathbb{E} \int_0^T |Z_s|^2 ds < +\infty \right\}$$

with $\mathbb{H}^{0,n}$ standing for the collection of all \mathbb{R}^n -valued progressively measurable processes on $[0, T]$. By (A.1) and (A.2), any $\alpha \in \mathbb{A}$ satisfies

$$\mathbb{E} \int_0^T [|b(t, 0, \delta_0, \alpha_t)|^2 + |\sigma(t, 0, \delta_0, \alpha_t)|^2] dt < +\infty.$$

Together with the Lipschitz assumption (A.2), this guarantees that, for any $\alpha \in \mathbb{A}$, there exists a unique solution $X = X^\alpha$ of (1), and that moreover this solution satisfies

$$(2) \quad \mathbb{E} \sup_{0 \leq t \leq T} |X_t|^p < +\infty$$

for every $p \in [1, 2]$. See e.g. [16, 10] for a proof.

The stochastic optimization problem which we consider is to minimize the objective function

$$(3) \quad J(\alpha) = \mathbb{E} \left\{ \int_0^T f(t, X_t, \mathbb{P}_{X_t}, \alpha_t) dt + g(X_T, \mathbb{P}_{X_T}) \right\},$$

over the set \mathbb{A} of admissible control processes $\alpha = (\alpha_t)_{0 \leq t \leq T}$. The *running cost* function f is a real valued deterministic function on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, and the *terminal cost* function g is a real valued deterministic function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. Assumptions on the cost functions f and g will be spelled out later.

The McKean-Vlasov dynamics posited in (1) are sometimes called of *mean field type*. This is justified by the fact that the uncontrolled stochastic differential equations of the McKean-Vlasov type first appeared in the infinite particle limit of large systems of particles with mean field interactions. See [12, 16, 10] for example. Typically, the dynamics of such a system of N particles are given by a system of N stochastic differential equations of the form

$$dX_t^i = b^i(t, X_t^1, \dots, X_t^N) dt + \sigma^i(t, X_t^1, \dots, X_t^N) dW_t^i,$$

where the W^i 's are N independent standard Wiener processes in \mathbb{R}^m , the σ^i 's are N deterministic functions from $[0, T] \times \mathbb{R}^{N \times d}$ into the space of $d \times m$ real matrices, and the b^i 's are N deterministic functions from $[0, T] \times \mathbb{R}^{N \times d}$ into \mathbb{R}^d . The interaction between the particles is said to be of mean field type when the functions b^i and σ^i are of the form

$$b^i(t, x_1, \dots, x_N) = b(t, x_i, \bar{\mu}_{\underline{x}}^N), \quad \text{and} \quad \sigma^i(t, x_1, \dots, x_N) = \sigma(t, x_i, \bar{\mu}_{\underline{x}}^N), \quad i = 1, \dots, N,$$

for some deterministic function b from $[0, T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$ into \mathbb{R}^d , and σ from $[0, T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d)$ into the space of $d \times m$ real matrices. Here, for each N -tuple $\underline{x} = (x_1, \dots, x_N)$, we denote by $\bar{\mu}_{\underline{x}}^N$, or $\bar{\mu}^N$ when no confusion is possible, the empirical probability measure defined by:

$$(4) \quad \bar{\mu}^N(dx') = \frac{1}{N} \sum_{j=1}^N \delta_{x_j}(dx')$$

and for each x by δ_x the unit point mass (Dirac) measure at x . We shall come back to this formulation of the problem in the last section of the paper when we use results from the propagation of chaos to construct approximate equilibriums.

We emphasize that the optimization problem (3) differs from the optimization problem encountered in the theory of mean field games. Differences between these optimization problems are discussed in [8]. When handling mean field games, the optimization of the cost functional (3) is performed for a fixed flow of probability measures. In other words, the argument $(\mathbb{P}_{X_t})_{0 \leq t \leq T}$ in (1) and (3) is kept fixed as α varies and the controlled processes are driven by the same flow of measures, which is not necessarily the flow of distributions of the process $(X_t)_{0 \leq t \leq T}$ but only an input. Solving the corresponding mean field game then consists in identifying a flow of probability measures, that is an input, such that the optimal states have precisely the input as flow of statistical distributions. As highlighted in Section 6, the optimization problem for controlled McKean-Vlasov dynamics as we consider here, also reads as a limit problem as N tends to infinity, of the optimal states of N interacting players or agents using a common policy.

Useful Notations. Given a function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ and a vector $p \in \mathbb{R}^d$, we will denote by $\partial h(x) \cdot p$ the action of the gradient of h onto p . When $h : \mathbb{R}^d \rightarrow \mathbb{R}^\ell$, we will also denote by $\partial h(x) \cdot p$ the action of the gradient of h onto p , the resulting quantity being an element of \mathbb{R}^ℓ . When $h : \mathbb{R}^d \rightarrow \mathbb{R}^\ell$ and $p \in \mathbb{R}^\ell$, we will denote by $\partial h(x) \odot p$ the element of \mathbb{R}^d defined by $\partial_x[h(x) \cdot p]$ where \cdot is here understood as the inner product in \mathbb{R}^ℓ .

3. PRELIMINARIES

We now introduce the notation and concepts needed for the analysis of the stochastic optimization problem associated with the control of McKean-Vlasov dynamics.

3.1. Differentiability and Convexity of Functions of Measures. There are many notions of differentiability for functions defined on spaces of measures, and recent progress in the theory of optimal transportation have put several forms in the limelight. See for example [1, 17] for exposés of these geometric approaches in textbook form. However, the notion of differentiability which we find convenient for the type of stochastic control problem studied in this paper is slightly different. It is more of a functional analytic nature. We believe that it was introduced by P.L. Lions in his lectures at the *Collège de France*. See [5] for a readable account. This notion of differentiability is based on the *lifting* of functions $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto H(\mu)$ into functions \tilde{H} defined on the Hilbert space $L^2(\tilde{\Omega}; \mathbb{R}^d)$ over some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ by setting $\tilde{H}(\tilde{X}) = H(\tilde{\mathbb{P}}_{\tilde{X}})$ for $\tilde{X} \in L^2(\tilde{\Omega}; \mathbb{R}^d)$, $\tilde{\Omega}$ being a Polish space and $\tilde{\mathbb{P}}$ an atomless measure. Then, a function H is said to be differentiable at $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ if there exists a random variable \tilde{X}_0 with law μ_0 , in other words satisfying $\tilde{\mathbb{P}}_{\tilde{X}_0} = \mu_0$, such that the lifted function \tilde{H} is Fréchet differentiable at \tilde{X}_0 . Whenever this is the case, the Fréchet derivative of \tilde{H} at \tilde{X}_0 can be viewed as an element of $L^2(\tilde{\Omega}; \mathbb{R}^d)$ by identifying $L^2(\tilde{\Omega}; \mathbb{R}^d)$ and its dual. It then turns out that its distribution depends only upon the law μ_0 and not upon the particular random variable \tilde{X}_0 having distribution μ_0 . See Section 6 in [5] for details. This Fréchet derivative $[D\tilde{H}](\tilde{X}_0)$ is called the representation of the derivative of H at μ_0 along the variable \tilde{X}_0 . Since it is viewed as an element of $L^2(\tilde{\Omega}; \mathbb{R}^d)$, by definition,

$$(5) \quad H(\mu) = H(\mu_0) + [D\tilde{H}](\tilde{X}_0) \cdot (\tilde{X} - \tilde{X}_0) + o(\|\tilde{X} - \tilde{X}_0\|_2)$$

whenever \tilde{X} and \tilde{X}_0 are random variables with distributions μ and μ_0 respectively, the dot product being here the L^2 - inner product over $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and $\|\cdot\|_2$ the associated norm. It is shown in [5] that, as a random variable, it is of the form $\tilde{h}(\tilde{X}_0)$ for some deterministic measurable function $\tilde{h} : \mathbb{R}^d \rightarrow \mathbb{R}^d$, which is uniquely defined μ_0 -almost everywhere on \mathbb{R}^d . The equivalence class of \tilde{h} in $L^2(\mathbb{R}^d, \mu_0)$ being uniquely defined, we can denote it by $\partial_\mu H(\mu_0)$ (or $\partial H(\mu_0)$ when no confusion is possible): We will call $\partial_\mu H(\mu_0)$ the derivative of H at μ_0 and we will often identify it with a function $\partial_\mu H(\mu_0)(\cdot) : \mathbb{R}^d \ni x \mapsto \partial_\mu H(\mu_0)(x) \in \mathbb{R}^d$ (or by $\partial H(\mu_0)(\cdot)$ when no confusion is possible). Notice that $\partial_\mu H(\mu_0)$ allows us to express $[D\tilde{H}](\tilde{X}_0)$ as a function of any random variable \tilde{X}_0 with distribution μ_0 , irrespective of where this random variable is defined. In particular, the differentiation formula (5) is somehow invariant by modification of the space $\tilde{\Omega}$ and of the variables \tilde{X}_0 and \tilde{X} used for the representation of H , in the sense that $[D\tilde{H}](\tilde{X}_0)$ always reads as $\partial_\mu H(\mu_0)(\tilde{X}_0)$, whatever the choices of $\tilde{\Omega}$, \tilde{X}_0 and \tilde{X} are. It is plain to see how this works when the function H is of the form

$$(6) \quad H(\mu) = \int_{\mathbb{R}^d} h(x) \mu(dx) = \langle h, \mu \rangle$$

for some scalar differentiable function h defined on \mathbb{R}^d . Indeed, in this case, $\tilde{H}(\tilde{X}) = \tilde{\mathbb{E}}[h(\tilde{X})]$ and $D\tilde{H}(\tilde{X}) \cdot \tilde{Y} = \tilde{\mathbb{E}}[\partial h(\tilde{X}) \cdot \tilde{Y}]$ so that we can think of $\partial_\mu H(\mu)$ as the deterministic function ∂h . We will use this particular example to recover the Pontryagin principle originally derived in [2] for scalar interactions as a particular case of the general Pontryagin principle which we prove below. The example (6) highlights the fact that this notion of differentiability is very different from the usual one. Indeed, given the fact that the function H defined by (6) is linear in the measure μ when viewed as an element of the dual of a function space, one should expect the derivative to be h and NOT h' ! The notion of differentiability used in this paper is best understood as differentiation of functions of limits of empirical measures (or linear combinations of Dirac point masses) in the directions of the atoms of the measures. We illustrate this fact in the next two propositions.

Proposition 3.1. *If u is differentiable on $\mathcal{P}_2(\mathbb{R}^d)$, given any integer $N \geq 1$, we define the empirical projection of u onto \mathbb{R}^d by*

$$\bar{u}^N : (\mathbb{R}^d)^N \ni \underline{x} = (x_1, \dots, x_N) \mapsto u\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right).$$

Then, \bar{u}^N is differentiable on $(\mathbb{R}^d)^N$ and, for all $i \in \{1, \dots, N\}$,

$$\partial_{x_i} \bar{u}^N(\underline{x}) = \partial_{x_i} \bar{u}^N(x_1, \dots, x_N) = \frac{1}{N} \partial u\left(\frac{1}{N} \sum_{j=1}^N \delta_{x_j}\right)(x_i).$$

Proof. On $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, consider a uniformly distributed random variable ϑ over the set $\{1, \dots, N\}$. Then, for any fixed $\underline{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$, x_ϑ is a random variable having the distribution $\bar{\mu}^N = N^{-1} \sum_{i=1}^N \delta_{x_i}$. In particular, with the same notation as above for \tilde{u} ,

$$\bar{u}^N(\underline{x}) = \bar{u}^N(x_1, \dots, x_N) = \tilde{u}(x_\vartheta).$$

Therefore, for $\underline{h} = (h_1, \dots, h_N) \in (\mathbb{R}^d)^N$,

$$\bar{u}^N(\underline{x} + \underline{h}) = \tilde{u}(x_\vartheta + h_\vartheta) = \tilde{u}(x_\vartheta) + D\tilde{u}(x_\vartheta) \cdot h_\vartheta + o(|h|),$$

the dot product being here the L^2 - inner product over $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, from which we deduce

$$\bar{u}^N(\underline{x} + \underline{h}) = \bar{u}^N(\underline{x}) + \frac{1}{N} \sum_{i=1}^N \partial u(\bar{\mu}^N)(x_i) h_i + o(|h|),$$

which is the desired result. \square

The mapping $D\tilde{u} : L^2(\tilde{\Omega}; \mathbb{R}^d) \rightarrow L^2(\tilde{\Omega}; \mathbb{R}^d)$ is said to be Lipschitz continuous if there exists a constant $C > 0$ such that, for any square integrable random variables \tilde{X} and \tilde{Y} in $L^2(\tilde{\Omega}; \mathbb{R}^d)$, it holds $\|D\tilde{u}(\tilde{X}) - D\tilde{u}(\tilde{Y})\|_2 \leq C\|\tilde{X} - \tilde{Y}\|_2$. In such a case, the Lipschitz property can be transferred onto $L^2(\Omega)$ and then rewritten as:

$$(7) \quad \mathbb{E}[|\partial u(\mathbb{P}_X)(X) - \partial u(\mathbb{P}_Y)(Y)|^2] \leq C^2 \mathbb{E}[|X - Y|^2],$$

for any square integrable random variables X and Y in $L^2(\Omega; \mathbb{R}^d)$. From our discussion of the construction of ∂u , notice that, for each μ , $\partial u(\mu)(\cdot)$ is only uniquely defined μ -almost everywhere. The following lemma (the proof of which is deferred to Subsection 3.3) then says that, in the current framework, there is a Lipschitz continuous version of $\partial u(\mu)(\cdot)$:

Lemma 3.2. *Given a family of Borel-measurable mappings $(v(\mu)(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d)_{\mu \in \mathcal{P}_2(\mathbb{R}^d)}$ indexed by the probability measures of order 2 on \mathbb{R}^d , assume that there exists a constant C such that, for any square integrable random variables ξ and ξ' in $L^2(\Omega; \mathbb{R}^d)$, it holds*

$$(8) \quad \mathbb{E}[|v(\mathbb{P}_\xi)(\xi) - v(\mathbb{P}_{\xi'})(\xi')|^2] \leq C^2 \mathbb{E}[|\xi - \xi'|^2].$$

Then, for each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, one can redefine $v(\mu)(\cdot)$ on a μ -negligeable set in such a way that:

$$\forall x, x' \in \mathbb{R}^d, \quad |v(\mu)(x) - v(\mu)(x')| \leq C|x - x'|,$$

for the same C as in (8).

By (7), we can use Lemma 3.2 in order to define $\partial u(\mu)(x)$ for every μ and every x while preserving the Lipschitz property in the variable x . From now on, we shall use this version of ∂u . So, if $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and X and Y are random variables such that $\mathbb{P}_X = \mu$ and $\mathbb{P}_Y = \nu$, we have:

$$\begin{aligned} \mathbb{E}[|\partial u(\mu)(X) - \partial u(\nu)(X)|^2] &\leq 2 \left(\mathbb{E}[|\partial u(\mu)(X) - \partial u(\nu)(Y)|^2] + \mathbb{E}[|\partial u(\nu)(Y) - \partial u(\nu)(X)|^2] \right) \\ &\leq 4C^2 \mathbb{E}[|Y - X|^2], \end{aligned}$$

where we used the Lipschitz property (7) of the derivative together with the result of Lemma 3.2 applied to the function $\partial u(\nu)$. Now, taking the infimum over all the couplings (X, Y) with marginals μ and ν , we obtain

$$\inf_{X, \mathbb{P}_X = \mu} \mathbb{E}[|\partial u(\mu)(X) - \partial u(\nu)(X)|^2] \leq 4C^2 W_2(\mathbb{P}_X, \mathbb{P}_Y)^2$$

and since the left-hand side depends only upon μ and not on X as long as $\mathbb{P}_X = \mu$, we get

$$(9) \quad \mathbb{E}[|\partial u(\mu)(X) - \partial u(\nu)(X)|^2] \leq 4C^2 W_2(\mu, \nu)^2.$$

We will use the following consequence of this estimate:

Proposition 3.3. *Let u be a differentiable function on $\mathcal{P}_2(\mathbb{R}^d)$ with a Lipschitz derivative, and let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and $\underline{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ and $\underline{y} = (y_1, \dots, y_N) \in (\mathbb{R}^d)^N$. Then, with the same notation as in the statement of Proposition 3.1, we have:*

$$\partial \bar{u}^N(\underline{x}) \cdot (\underline{y} - \underline{x}) = \frac{1}{N} \sum_{i=1}^N \partial u(\mu)(x_i)(y_i - x_i) + \mathcal{O} \left[W_2(\bar{\mu}^N, \mu) \left(N^{-1} \sum_{i=1}^N |x_i - y_i|^2 \right)^{1/2} \right],$$

the dot product being here the usual Euclidean inner product and \mathcal{O} standing for the Landau notation.

Proof. Using Proposition 3.1, we get:

$$\begin{aligned} \partial \bar{u}^N(\underline{x}) \cdot (\underline{y} - \underline{x}) &= \sum_{i=1}^N \partial_{x_i} \bar{u}^N(\underline{x})(y_i - x_i) \\ &= \frac{1}{N} \sum_{i=1}^N \partial u(\bar{\mu}^N)(x_i)(y_i - x_i) \\ &= \frac{1}{N} \sum_{i=1}^N \partial u(\mu)(x_i)(y_i - x_i) + \frac{1}{N} \sum_{i=1}^N [\partial u(\bar{\mu}^N)(x_i) - \partial u(\mu)(x_i)](y_i - x_i). \end{aligned}$$

Now, by Cauchy-Schwarz' inequality,

$$\begin{aligned} &\left| \frac{1}{N} \sum_{i=1}^N [\partial u(\bar{\mu}^N)(x_i) - \partial u(\mu)(x_i)](y_i - x_i) \right| \\ &\leq \left(\frac{1}{N} \sum_{i=1}^N |\partial u(\bar{\mu}^N)(x_i) - \partial u(\mu)(x_i)|^2 \right)^{1/2} \left(\frac{1}{N} \sum_{i=1}^N |y_i - x_i|^2 \right)^{1/2} \\ &= (\tilde{\mathbb{E}}[|\partial u(\bar{\mu}^N)(x_\vartheta) - \partial u(\mu)(x_\vartheta)|^2])^{1/2} \left(\frac{1}{N} \sum_{i=1}^N |y_i - x_i|^2 \right)^{1/2} \\ &\leq 2C W_2(\bar{\mu}^N, \mu) \left(\frac{1}{N} \sum_{i=1}^N |y_i - x_i|^2 \right)^{1/2}, \end{aligned}$$

if we use the same notation for ϑ as in the proof of Proposition 3.1 and apply the estimate (9) with $X = x_\vartheta$, $\mu = \bar{\mu}^N$ and $\nu = \mu$. \square

Remark 3.4. *We shall use the estimate of Proposition 3.3 when $x_i = X_i$ and the X_i 's are independent \mathbb{R}^d -valued random variables with common distribution μ . Whenever $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the law of large numbers ensures that the Wasserstein distance between μ and the empirical measure $\bar{\mu}^N$ tends to 0 a.s., that is*

$$\mathbb{P} \left(\lim_{n \rightarrow +\infty} W_2(\bar{\mu}^N, \mu) = 0 \right) = 1,$$

see for example Section 10 in [15]. Since we can find a constant $C > 0$, independent of N , such that

$$W_2^2(\bar{\mu}^N, \mu) \leq C \left(1 + \frac{1}{N} \sum_{i=1}^N |X_i|^2 \right),$$

we deduce from the law of large numbers again that the family $(W_2^2(\bar{\mu}^N, \mu))_{N \geq 1}$ is uniformly integrable, so that the convergence to 0 also holds in the L^2 sense:

$$(10) \quad \lim_{n \rightarrow +\infty} \mathbb{E}[W_2^2(\bar{\mu}^N, \mu)] = 0.$$

Whenever $\int_{\mathbb{R}^d} |x|^{d+5} \mu(dx) < \infty$, the rate of convergence can be specified. We indeed have the following standard estimate on the Wasserstein distance between μ and the empirical measure $\bar{\mu}^N$:

$$(11) \quad \mathbb{E}[W_2^2(\bar{\mu}^N, \mu)] \leq CN^{-2/(d+4)},$$

for some constant $C > 0$, see for example Section 10 in [15]. Proposition 3.3 then says that, when N is large, the gradient of \bar{u}^N at the empirical sample $(X_i)_{1 \leq i \leq N}$ is close to the sample $(\partial u(\mu)(X_i))_{1 \leq i \leq N}$, the accuracy of the approximation being specified in the $L^2(\Omega)$ norm by (11) when μ is sufficiently integrable.

3.2. Joint Differentiability and Convexity.

Joint Differentiability. Below, we often consider functions $g : \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, \mu) \rightarrow g(x, \mu) \in \mathbb{R}$ depending on both an n -dimensional x and a probability measure μ . Joint differentiability is then defined according to the same procedure: g is said to be jointly differentiable if the *lifting* $\tilde{g} : \mathbb{R}^n \times L^2(\tilde{\Omega}; \mathbb{R}^d) \ni (x, \tilde{X}) \mapsto g(x, \tilde{\mathbb{P}}_{\tilde{X}})$ is jointly differentiable. In such a case, we can define the partial derivatives in x and μ : they read $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, \mu) \mapsto \partial_x g(x, \mu)$ and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, \mu) \mapsto \partial_\mu g(x, \mu)(\cdot) \in L^2(\mathbb{R}^d, \mu)$ respectively. The partial Fréchet derivative of \tilde{g} in the direction \tilde{X} thus reads $L^2(\tilde{\Omega}; \mathbb{R}^d) \ni (x, \tilde{X}) \mapsto D_{\tilde{X}} \tilde{g}(x, \tilde{X}) = \partial_\mu g(x, \tilde{\mathbb{P}}_{\tilde{X}})(\tilde{X}) \in L^2(\tilde{\Omega}; \mathbb{R}^d)$.

We often use the fact that joint continuous differentiability in the two arguments is equivalent with partial differentiability in each of the two arguments and joint continuity of the partial derivatives. Here, the joint continuity of $\partial_x g$ is understood as the joint continuity with respect to the Euclidean distance on \mathbb{R}^n and the Wasserstein distance on $\mathcal{P}_2(\mathbb{R}^d)$. The joint continuity of $\partial_\mu g$ is understood as the joint continuity of the mapping $(x, \tilde{X}) \mapsto \partial_\mu g(x, \tilde{\mathbb{P}}_{\tilde{X}})(\tilde{X})$ from $\mathbb{R}^n \times L^2(\tilde{\Omega}; \mathbb{R}^d)$ into $L^2(\tilde{\Omega}; \mathbb{R}^d)$.

When the partial derivatives of g are assumed to be Lipschitz-continuous, we can benefit from Lemma 3.2. It says that, for any (x, μ) , the *representation* $\mathbb{R}^d \ni x' \mapsto \partial_\mu g(x, \mu)(x')$ makes sense as a Lipschitz function in x' and that an appropriate version of (9) holds true.

Convex Functions of Measures. We define a notion of convexity associated with this notion of differentiability. A function g on $\mathcal{P}_2(\mathbb{R}^d)$ which is differentiable in the above sense is said to be convex if for every μ and μ' in $\mathcal{P}_2(\mathbb{R}^d)$ we have

$$(12) \quad g(\mu') - g(\mu) - \tilde{\mathbb{E}}[\partial_\mu g(\mu)(\tilde{X}) \cdot (\tilde{X}' - \tilde{X})] \geq 0$$

whenever \tilde{X} and \tilde{X}' are square integrable random variables with distributions μ and μ' respectively. Examples are given in Subsection 4.3.

More generally, a function g on $\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^d)$ which is jointly differentiable in the above sense is said to be convex if for every (x, μ) and (x', μ') in $\mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^d)$ we have

$$(13) \quad g(x', \mu') - g(x, \mu) - \partial_x g(x, \mu) \cdot (x' - x) - \tilde{\mathbb{E}}[\partial_\mu g(x, \mu)(\tilde{X}) \cdot (\tilde{X}' - \tilde{X})] \geq 0$$

whenever \tilde{X} and \tilde{X}' are square integrable random variables with distributions μ and μ' respectively.

3.3. Proof of Lemma 3.2. First Step. We first consider the case v bounded and assume that μ has a strictly positive continuous density p on the whole \mathbb{R}^d , p and its derivatives being of exponential decay at the infinity. We then claim that there exists a continuously differentiable one-to-one function from $(0, 1)^d$ onto \mathbb{R}^d such that, whenever η_1, \dots, η_d are d independent random variables, each of them being uniformly distributed on $(0, 1)$, $U(\eta_1, \dots, \eta_d)$ has distribution μ . It satisfies for any $(z_1, \dots, z_d) \in (0, 1)^d$

$$\frac{\partial U_i}{\partial z_i}(z_1, \dots, z_d) \neq 0, \quad \frac{\partial U_j}{\partial z_i}(z_1, \dots, z_d) = 0, \quad 1 \leq i < j \leq d.$$

The result is well-known when $d = 1$. In such a case, U is the inverse of the cumulative distribution function of μ . In higher dimension, U can be constructed by an induction argument on the dimension. Assume indeed that some \hat{U} has been constructed for the first marginal distribution $\hat{\mu}$ of μ on \mathbb{R}^{d-1} , that is for the push-forward of μ by the projection mapping $\mathbb{R}^d \ni (x_1, \dots, x_d) \mapsto (x_1, \dots, x_{d-1})$. Given $(x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$, we then denote by $p(\cdot | x_1, \dots, x_{d-1})$ the conditional density of μ given the $d - 1$ first coordinates:

$$p(x_d | x_1, \dots, x_{d-1}) = \frac{p(x_1, \dots, x_d)}{\hat{p}(x_1, \dots, x_{d-1})}, \quad x_1, \dots, x_{d-1} \in \mathbb{R}^{d-1},$$

where \hat{p} denotes the density of $\hat{\mu}$ (which is continuously differentiable and positive). We then denote by $(0, 1) \ni z_d \mapsto U_d(z_d | x_1, \dots, x_{d-1})$ the inverse of the cumulative distribution function of the law of density $p(\cdot | x_1, \dots, x_{d-1})$. It satisfies

$$F_d(U_d(z_d | x_1, \dots, x_{d-1}) | x_1, \dots, x_{d-1}) = z_d,$$

with

$$F_d(x_d | x_1, \dots, x_{d-1}) = \int_{-\infty}^{x_d} p(y | x_1, \dots, x_{d-1}) dy,$$

which is continuously differentiable in (x_1, \dots, x_d) (using the exponential decay of the density at the infinity). By the implicit function theorem, the mapping $\mathbb{R}^{d-1} \times (0, 1) \ni (x_1, \dots, x_{d-1}, z_d) \mapsto U_d(z_d | x_1, \dots, x_{d-1})$ is continuously differentiable. The partial derivative with respect to z_d is given by

$$\frac{\partial U_d}{\partial z_d}(z_d | x_1, \dots, x_{d-1}) = \frac{1}{p(U_d(z_d | x_1, \dots, x_{d-1}) | x_1, \dots, x_{d-1})},$$

which is non-zero. We now let

$$U(z_1, \dots, z_d) = (\hat{U}(z_1, \dots, z_{d-1}), U_d(z_d | \hat{U}(z_1, \dots, z_{d-1}))), \quad z_1, \dots, z_d \in (0, 1)^d.$$

By construction, $U(\eta_1, \dots, \eta_d)$ has distribution μ : $(\eta_1, \dots, \eta_{d-1})$ has distribution $\hat{\mu}$ and the conditional law of $U_d(\eta_1, \dots, \eta_d)$ given $\eta_1, \dots, \eta_{d-1}$ is the conditional law of μ given the $d - 1$ first coordinates. It satisfies $[\partial U_d / \partial z_d](z_1, \dots, z_d) > 0$ and $[\partial U_i / \partial z_d](z_1, \dots, z_d) = 0$ for $i < d$. In particular, since \hat{U} is assumed (by induction) to be injective and $[\partial U_d / \partial z_d](z_1, \dots, z_d) > 0$, U must be injective as well. As the Jacobian matrix of U is triangular with non-zero elements on the diagonal, it is invertible. By the global inversion theorem, U is a diffeomorphism: the range of U is the support of μ , that is \mathbb{R}^d . This proves that U is one-to-one from $(0, 1)^d$ onto \mathbb{R}^d .

Second Step. We still consider the case v bounded and assume that μ has a strictly positive continuous density p on the whole \mathbb{R}^d , p and its derivatives being of exponential decay at the infinity. We will use the mapping U constructed in the first step. For three random variables ξ , ξ' and G in

$L^2(\Omega; \mathbb{R}^d)$, the pair (ξ, ξ') being independent of G , the random variables ξ and ξ' having the same distribution, and G being $\mathcal{N}_d(0, I_d)$ normally distributed, (8) implies that, for any integer $n \geq 1$

$$\mathbb{E}[|v(\xi + n^{-1}G, \mathbb{P}_{\xi+n^{-1}G}) - v(\xi' + n^{-1}G, \mathbb{P}_{\xi+n^{-1}G})|^2] \leq C^2 \mathbb{E}[|\xi - \xi'|^2].$$

In particular, setting

$$v_n(x) = \frac{n^d}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} v(y, \mathbb{P}_{\xi+n^{-1}G}) \exp(-n^2 \frac{|x-y|^2}{2}) dy,$$

we have

$$(14) \quad \mathbb{E}[|v_n(\xi) - v_n(\xi')|^2] \leq C^2 \mathbb{E}[|\xi - \xi'|^2].$$

Notice that v_n is infinitely differentiable with bounded derivatives.

We now choose a specific coupling for ξ and ξ' . Indeed, we know that for any $\eta = (\eta_1, \dots, \eta_d)$ and $\eta' = (\eta'_1, \dots, \eta'_d)$, with uniform distribution on $(0, 1)^d$, $U(\eta)$ and $U(\eta')$ have the same distribution as ξ . Without any loss of generality, we then assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is given by $(0, 1)^d \times \mathbb{R}^d$ endowed with its Borel σ -algebra and the product of the Lebesgue measure on $(0, 1)^d$ and of the Gaussian measure $\mathcal{N}_d(0, I_d)$. The random variables η and G are then chosen as the canonical mappings $\eta : (0, 1)^d \times \mathbb{R}^d \ni (z, y) \mapsto z$ and $G : (0, 1)^d \times \mathbb{R}^d \ni (z, y) \mapsto y$.

We then define η' as a function of the variable $z \in (0, 1)^d$ only. For a given $z^0 = (z_1^0, \dots, z_d^0) \in (0, 1)^d$ and for h small enough so that the open ball $B(z^0, h)$ of center z^0 and radius h is included in $(0, 1)^d$, we let:

$$\eta'(z) = \begin{cases} z^0 + (z_d^0 - z_d)e_d & \text{on } B(z^0, h), \\ z, & \text{outside } B(z^0, h), \end{cases}$$

where e_d is the d th vector of the canonical basis. We rewrite (14) as:

$$\int_{(0,1)^d} |v_n(U(\eta(z))) - v_n(U(\eta'(z)))|^2 dz \leq C^2 \int_{(0,1)^d} |U(\eta(z)) - U(\eta'(z))|^2 dz,$$

or equivalently:

$$(15) \quad \begin{aligned} & \int_{|r|<h} |v_n[U(z^0 + r - 2r_de_d)] - v_n(U(z^0 + r))|^2 dr \\ & \leq C^2 \int_{|r|<h} |U(z^0 + r - 2r_de_d) - U(z^0 + r)|^2 dr. \end{aligned}$$

Since U is continuously differentiable, we have

$$v_n(U(z^0 + r)) = v_n(U(z^0)) + \partial v_n(U(z^0)) \cdot [\partial U(z^0) \cdot r] + o(r),$$

where $\partial U(z^0)$ is a $d \times d$ matrix. We deduce

$$\begin{aligned} v_n[U(z^0 + r - 2r_de_d)] - v_n(U(z^0 + r)) &= -2 \sum_{i=1}^d \frac{\partial v_n}{\partial x_i}(U(z^0)) \frac{\partial U_i}{\partial z_d}(z^0) r_d + o(r) \\ &= -2 \frac{\partial v_n}{\partial x_d}(U(z^0)) \frac{\partial U_d}{\partial z_d}(z^0) r_d + o(r), \end{aligned}$$

since $\partial U_i / \partial z_d = 0$ for $i \neq d$, and

$$(16) \quad \begin{aligned} & \int_{|r|<h} |v_n[U(z^0 + r - 2r_d e_d)] - v_n(U(z^0 - r))|^2 dr \\ &= 4 \left| \frac{\partial v_n}{\partial x_d}(U(z^0)) \frac{\partial U_d}{\partial z_d}(z^0) \right|^2 \int_{|r|<h} r_d^2 dr + o(h^3). \end{aligned}$$

Similarly,

$$(17) \quad \int_{|r|<h} |U(z^0 + r - 2r_d e_d) - U(z^0 + r)|^2 dr = 4 \left| \frac{\partial U_d}{\partial z_d}(z^0) \right|^2 \int_{|r|<h} r_d^2 dr + o(h^3).$$

and putting together (15), (16), and (17), we obtain

$$\left| \frac{\partial v_n}{\partial x_d}(U(z^0)) \frac{\partial U_d}{\partial z_d}(z^0) \right|^2 \leq C^2 \left| \frac{\partial U_d}{\partial z_d}(z^0) \right|^2.$$

Since $[\partial U_d / \partial z_d](z^0)$ is different from zero, we deduce that

$$\left| \frac{\partial v_n}{\partial x_d}(U(z^0)) \right|^2 \leq C^2,$$

and since U is a one-to-one mapping from $(0, 1)^d$ onto \mathbb{R}^d , and $z^0 \in (0, 1)^d$ is arbitrary, we conclude that $|\partial v_n / \partial x_d|(x)| \leq C$, for any $x \in \mathbb{R}^d$. By changing the basis used for the construction of U (we used the canonical one but we could use any orthonormal basis), we have $|\nabla v_n(x)e| \leq C$ for any $x, e \in \mathbb{R}^d$ with $|e| = 1$. This proves that the functions $(v_n)_{n \geq 1}$ are uniformly bounded and C -Lipschitz continuous. We then denote by \hat{v} the limit of a subsequence converging for the topology of uniform convergence on compact subsets. For simplicity, we keep the index n to denote the subsequence. Assumption (8) implies:

$$\mathbb{E}[|v_n(\xi) - v(\xi, \mathbb{P}_\xi)|^2] \leq \mathbb{E}[|v(\xi + n^{-1}G, \mathbb{P}_{\xi+n^{-1}G}) - v(\xi, \mathbb{P}_\xi)|^2] \leq C^2 n^{-2},$$

and taking the limit $n \rightarrow +\infty$, we deduce that \hat{v} and $v(\cdot, \mathbb{P}_\xi)$ coincide \mathbb{P}_ξ almost everywhere. This completes the proof when v is bounded and ξ has a continuous positive density p , p and its derivatives being of exponential decay at the infinity.

Third Step. When v is bounded and ξ is bounded and has a general distribution, we approximate ξ by $\xi + n^{-1}G$ again. Then, $\xi + n^{-1}G$ has a positive continuous density, the density and its derivatives being of Gaussian decay at the infinity, so that, by the second step, the function $\mathbb{R}^d \ni x \mapsto v(x, \mathbb{P}_{\xi+n^{-1}G})$ can be assumed to be C -Lipschitz continuous for each $n \geq 1$. Extracting a converging subsequence and passing to the limit as above, we deduce that $v(\cdot, \mathbb{P}_\xi)$ admits a C -Lipschitz continuous version.

When v is bounded but ξ is not bounded, we approximate ξ by its orthogonal projection on the ball of center 0 and radius n . We then complete the proof in a similar way.

Finally when v is not bounded, we approximate v by $(\psi_n(v))_{n \geq 1}$ where, for each $n \geq 1$, ψ_n is a bounded smooth function from \mathbb{R} into itself such that $\psi_n(r) = r$ for $r \in [-n, n]$ and $|\psi_n(r)| \leq 1$ for all $r \in \mathbb{R}$. Then, for each $n \geq 1$, there exists a C -Lipschitz continuous version of $\psi_n(v(\cdot, \mathbb{P}_\xi))$. Choosing some $x_0 \in \mathbb{R}^d$ such that $|v(x_0, \mathbb{P}_\xi)| < +\infty$, the sequence $\psi_n(v(x_0, \mathbb{P}_\xi))$ is bounded so that the sequence of functions $(\psi_n(v(\cdot, \mathbb{P}_\xi)))_{n \geq 1}$ is uniformly bounded and continuous on compact subsets. Extracting a converging subsequence, we complete the proof in the same way as before.

3.4. The Hamiltonian and the Dual Equations. The Hamiltonian of the stochastic optimization problem is defined as the function H given by

$$(18) \quad H(t, x, \mu, y, z, \alpha) = b(t, x, \mu, \alpha) \cdot y + \sigma(t, x, \mu, \alpha) \cdot z + f(t, x, \mu, \alpha)$$

where the dot notation stands here for the inner product in an Euclidean space. Because we need to compute derivatives of H with respect to its variable μ , we consider the lifting \tilde{H} defined by

$$(19) \quad \tilde{H}(t, x, \tilde{X}, y, z, \alpha) = H(t, x, \mu, y, z, \alpha)$$

for any random variable \tilde{X} with distribution μ , and we shall denote by $\partial_\mu H(t, x, \mu_0, y, z, \alpha)$ the derivative with respect to μ computed at μ_0 (as defined above) whenever all the other variables t, x, y, z and α are held fixed. We recall that $\partial_\mu H(t, x, \mu_0, y, z, \alpha)$ is an element of $L^2(\mathbb{R}^d, \mu_0)$ and that we identify it with a function $\partial_\mu H(t, x, \mu_0, y, z, \alpha)(\cdot) : \mathbb{R}^d \ni \tilde{x} \mapsto \partial_\mu H(t, x, \mu_0, y, z, \alpha)(\tilde{x})$. It satisfies $D\tilde{H}(t, x, \tilde{X}, y, z, \alpha) = \partial_\mu H(t, x, \mu_0, y, z, \alpha)(\tilde{X})$ almost-surely under $\tilde{\mathbb{P}}$.

Definition 3.5. *In addition to (A1-2), assume that the coefficients b, σ, f and g are differentiable with respect to x and μ . Then, given an admissible control $\alpha = (\alpha_t)_{0 \leq t \leq T} \in \mathbb{A}$, we denote by $X = X^\alpha$ the corresponding controlled state process. Whenever*

$$(20) \quad \mathbb{E} \int_0^T \{ |\partial_x f(t, X_t, \mathbb{P}_{X_t}, \alpha_t)|^2 + \tilde{\mathbb{E}}[|\partial_\mu f(t, X_t, \mathbb{P}_{X_t}, \alpha_t)(\tilde{X}_t)|^2] \} dt < +\infty,$$

and

$$(21) \quad \mathbb{E}\{ |\partial_x g(X_T, \mathbb{P}_{X_T})|^2 + \tilde{\mathbb{E}}[|\partial_\mu g(X_T, \mathbb{P}_{X_T})(\tilde{X}_T)|^2] \} < +\infty,$$

we call adjoint processes of X any couple $((Y_t)_{0 \leq t \leq T}, (Z_t)_{0 \leq t \leq T})$ of progressively measurable stochastic processes in $\mathbb{H}^{2,d} \times \mathbb{H}^{2,d \times m}$ satisfying the equation (which we call the adjoint equation):

$$(22) \quad \begin{cases} dY_t = -\partial_x H(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t) dt + Z_t dW_t \\ \quad - \tilde{\mathbb{E}}[\partial_\mu \tilde{H}(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)(X_t)] dt \\ Y_T = \partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)] \end{cases}$$

where $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{\alpha})$ is an independent copy of (X, Y, Z, α) defined on $L^2(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and $\tilde{\mathbb{E}}$ denotes the expectation on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Notice that $\mathbb{E}[\partial_\mu \tilde{H}(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)(X_t)]$ is a function of the random variable X_t as it stands for $\mathbb{E}[\partial_\mu \tilde{H}(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)(x)]|_{x=X_t}$ (and similarly for $\tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]$). Notice that, when b, σ, f and g do not depend upon the marginal distributions of the controlled state process, the extra terms appearing in the adjoint equation and its terminal condition disappear and this equation coincides with the classical adjoint equation of stochastic control.

Using the appropriate interpretation of the symbol \odot as explained in Section 2 and extending this notation to derivatives of the form $\partial_\mu h(\mu)(x) \odot p = (\partial_\mu [h(\mu) \cdot p])(x)$, the adjoint equation rewrites

$$(23) \quad \begin{aligned} dY_t = & -[\partial_x b(t, X_t, \mathbb{P}_{X_t}, \alpha_t) \odot Y_t + \partial_x \sigma(t, X_t, \mathbb{P}_{X_t}, \alpha_t) \odot Z_t + \partial_x f(t, X_t, \mathbb{P}_{X_t}, \alpha_t)] dt \\ & + Z_t dW_t \\ & - \tilde{\mathbb{E}}[\partial_\mu b(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Y}_t + \partial_\mu \sigma(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Z}_t \\ & + \partial_\mu f(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t)] dt, \end{aligned}$$

with the terminal condition $Y_T = \partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]$. Notice that $\partial_x b$ and $\partial_x \sigma$ are bounded since b and σ are assumed to be c -Lipschitz continuous in the variable x , see (A2). Notice also that $\mathbb{E}[|\partial_\mu b(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t)|^2]^{1/2}$ and $\mathbb{E}[|\partial_\mu \sigma(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t)|^2]^{1/2}$ are also bounded by c since b and σ are assumed to be c -Lipschitz continuous in the variable μ with respect to the 2-Wasserstein distance. It is indeed plain to check that, given a differentiable function $h : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, the notion of differentiability being defined as above, it holds $\mathbb{E}[|\partial_\mu h(X)|^2]^{1/2} \leq c$, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and any random variable X having μ as distribution, when h is c -Lipschitz continuous in μ with respect to the 2-Wasserstein distance.

Notice finally that, given an admissible control $\alpha \in \mathbb{A}$ and the corresponding controlled state process $X = X^\alpha$, despite the conditions (20–21) and despite the fact that the first part of the equation appears to be linear in the unknown processes Y_t and Z_t , existence and uniqueness of a solution (Y, Z) of the adjoint equation is not provided by standard results on Backward Stochastic Differential Equations (BSDEs) as the distributions of the solution processes (more precisely their joint distributions with the control and state processes α and X) appear in the coefficients of the equation. However, a slight modification of the original existence and uniqueness result of Pardoux and Peng [13] shows that existence and uniqueness still hold in our more general setting. The main lines of the proof are given in [4], Proposition 3.1 and Lemma 3.1. However, Lemma 3.1 in [4] doesn't apply directly since the coefficients $(\partial_\mu b(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Y}_t)_{0 \leq t \leq T}$ and $(\partial_\mu \sigma(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Z}_t)_{0 \leq t \leq T}$ are not Lipschitz continuous in \tilde{Y} and \tilde{Z} uniformly in the randomness, see Condition (C1) in [4]. Actually, a careful inspection of the proof shows that the bounds

$$\begin{aligned} \mathbb{E}\tilde{\mathbb{E}}[|\partial_\mu b(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Y}_t|^2] &\leq c'\mathbb{E}[|Y_t|^2], \\ \mathbb{E}\tilde{\mathbb{E}}[|\partial_\mu \sigma(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t) \odot \tilde{Z}_t|^2] &\leq c'\mathbb{E}[|Z_t|^2], \end{aligned}$$

are sufficient to make the whole argument work and thus to prove existence and uniqueness of a solution (Y, Z) satisfying

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t|^2 + \int_0^T |Z_t|^2 dt\right] < +\infty.$$

4. PONTRYAGIN PRINCIPLE FOR OPTIMALITY

In this section, we discuss sufficient and necessary conditions for optimality when the Hamiltonian satisfies appropriate assumptions of convexity. These conditions will be specified next, depending upon the framework. For the time being, we detail the regularity properties that will be used throughout the section. Referring to Subsection 3.2 for definitions of joint differentiability, we assume:

(A3) The functions b , σ and f are differentiable with respect to (x, α) , the mappings $(x, \mu, \alpha) \mapsto \partial_x(b, \sigma, f)(t, x, \mu, \alpha)$ and $(x, \mu, \alpha) \mapsto \partial_\alpha(b, \sigma, f)(t, x, \mu, \alpha)$ being continuous for any $t \in [0, T]$. The functions b , σ and f are also differentiable with respect to the variable μ in the sense given above, the mapping $\mathbb{R}^d \times L^2(\Omega; \mathbb{R}^d) \times A \ni (x, X, \alpha) \mapsto \partial_\mu(b, \sigma, f)(t, x, \mathbb{P}_X, \alpha)(X) \in L^2(\Omega; \mathbb{R}^{d \times d} \times \mathbb{R}^{(d \times m) \times d} \times \mathbb{R}^d)$ being continuous for any $t \in [0, T]$. Similarly, the function g is differentiable with respect to x , the mapping $(x, \mu) \mapsto \partial_x g(x, \mu)$ being continuous. The function g is also differentiable with respect to the variable μ , the mapping $\mathbb{R}^d \times L^2(\Omega; \mathbb{R}^d) \mapsto \partial_\mu g(x, \mathbb{P}_X)(X) \in L^2(\Omega; \mathbb{R}^d)$ being continuous.

(A4) The coefficients $((b, \sigma, f)(t, 0, \delta_0, 0))_{0 \leq t \leq T}$ are uniformly bounded. The partial derivatives $\partial_x b$ and $\partial_\alpha b$ are uniformly bounded and the norm of the mapping $x' \mapsto \partial_\mu(b, \sigma)(t, x, \mu, \alpha)(x')$ in $L^2(\mathbb{R}^d, \mu)$ is also uniformly bounded (i.e. uniformly in (t, x, μ, α)). There exists a constant L such that, for any $R \geq 0$ and any (t, x, μ, α) such that $|x|, \|\mu\|_2, |\alpha| \leq R$, $|\partial_x(f, g)(t, x, \mu, \alpha)|$ and $|\partial_\alpha f(t, x, \mu, \alpha)|$ are bounded by $L(1 + R)$ and the $L^2(\mathbb{R}^d, \mu)$ -norm of $x' \mapsto \partial_\mu(f, g)(t, x, \mu, \alpha)(x')$ is bounded by $L(1 + R)$. Here, we have used the notation

$$\|\mu\|_2^2 = \int_{\mathbb{R}^d} |x|^2 d\mu(x), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

4.1. A Necessary Condition. We assume that the sets A and \mathbb{A} of admissible controls are convex, we fix $\alpha \in \mathbb{A}$, and as before, we denote by $X = X^\alpha$ the corresponding controlled state process, namely the solution of (1) with given initial condition $X_0 = x_0$. Our first task is to compute the Gâteaux derivative of the cost functional J at α in all directions. In order to do so, we choose $\beta \in \mathbb{H}^{2,k}$ such that $\alpha + \epsilon\beta \in \mathbb{A}$ for $\epsilon > 0$ small enough. We then compute the variation of J at α in the direction of β (think of β as the difference between another element of \mathbb{A} and α).

Letting $(\theta_t = (X_t, \mathbb{P}_{X_t}, \alpha_t))_{0 \leq t \leq T}$, we define the variation process $V = (V_t)_{0 \leq t \leq T}$ to be the solution of the equation

$$(24) \quad dV_t = [\gamma_t \cdot V_t + \delta_t(\mathbb{P}_{(X_t, V_t)}) + \eta_t] dt + [\tilde{\gamma}_t \cdot V_t + \tilde{\delta}_t(\mathbb{P}_{(X_t, V_t)}) + \tilde{\eta}_t] dW_t,$$

with $V_0 = 0$, where the coefficients $\gamma_t, \delta_t, \eta_t, \tilde{\gamma}_t, \tilde{\delta}_t$ and $\tilde{\eta}_t$ are defined as

$$\gamma_t = \partial_x b(t, \theta_t), \quad \tilde{\gamma}_t = \partial_x \sigma(t, \theta_t), \quad \eta_t = \partial_\alpha b(t, \theta_t) \cdot \beta_t, \quad \text{and} \quad \tilde{\eta}_t = \partial_\alpha \sigma(t, \theta_t) \cdot \beta_t,$$

which are progressively measurable bounded processes with values in $\mathbb{R}^{d \times d}$, $\mathbb{R}^{(d \times m) \times d}$, \mathbb{R}^d , and $\mathbb{R}^{d \times m}$ respectively (the parentheses around $d \times m$ indicating that $\tilde{\gamma}_t \cdot u$ is seen as an element of $\mathbb{R}^{d \times m}$ whenever $u \in \mathbb{R}^d$), and

$$(25) \quad \begin{aligned} \delta_t &= \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] = \tilde{\mathbb{E}}[\partial_\mu b(t, x, \mathbb{P}_{X_t}, \alpha)(\tilde{X}_t) \cdot \tilde{V}_t] \Big|_{\substack{x=X_t \\ \alpha=\alpha_t}}, \\ \tilde{\delta}_t &= \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] = \tilde{\mathbb{E}}[\partial_\mu \sigma(t, x, \mathbb{P}_{X_t}, \alpha)(\tilde{X}_t) \cdot \tilde{V}_t] \Big|_{\substack{x=X_t \\ \alpha=\alpha_t}}, \end{aligned}$$

which are progressively measurable bounded processes with values in \mathbb{R}^d and $\mathbb{R}^{d \times m}$ respectively and where $(\tilde{X}_t, \tilde{V}_t)$ is an independent copy of (X_t, V_t) . As expectations of functions of $(\tilde{X}_t, \tilde{V}_t)$, δ_t and $\tilde{\delta}_t$ depend upon the joint distribution of X_t and V_t . In (24) we wrote $\delta_t(\mathbb{P}_{(X_t, V_t)})$ and $\tilde{\delta}_t(\mathbb{P}_{(X_t, V_t)})$ in order to stress the dependence upon the joint distribution of X_t and V_t . Even though we are dealing with possibly random coefficients, the existence and uniqueness of the variation process is guaranteed by Proposition 2.1 of [10] applied to the couple (X, V) and the system formed by (1) and (24). Because of our assumption on the boundedness of the partial derivatives of the coefficients, V satisfies $\mathbb{E} \sup_{0 \leq t \leq T} |V_t|^p < \infty$ for every finite $p \geq 1$.

Lemma 4.1. *For each $\epsilon > 0$ small enough, we denote by α^ϵ the admissible control defined by $\alpha_t^\epsilon = \alpha_t + \epsilon\beta_t$, and by $X^\epsilon = X^{\alpha^\epsilon}$ the corresponding controlled state. We have:*

$$(26) \quad \lim_{\epsilon \searrow 0} \mathbb{E} \sup_{0 \leq t \leq T} \left| \frac{X_t^\epsilon - X_t}{\epsilon} - V_t \right|^2 = 0.$$

Proof. For the purpose of this proof we set $\theta_t^\epsilon = (X_t^\epsilon, \mathbb{P}_{X_t^\epsilon}, \alpha_t^\epsilon)$ and $V_t^\epsilon = \epsilon^{-1}(X_t^\epsilon - X_t) - V_t$. Notice that $V_0^\epsilon = 0$ and that

$$\begin{aligned}
 dV_t^\epsilon &= \left[\frac{1}{\epsilon} [b(t, \theta_t^\epsilon) - b(t, \theta_t)] - \partial_x b(t, \theta_t) \cdot V_t - \partial_\alpha b(t, \theta_t) \cdot \beta_t - \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] \right] dt \\
 (27) \quad &+ \left[\frac{1}{\epsilon} [\sigma(t, \theta_t^\epsilon) - \sigma(t, \theta_t)] - \partial_x \sigma(t, \theta_t) \cdot V_t - \partial_\alpha \sigma(t, \theta_t) \cdot \beta_t - \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] \right] dW_t \\
 &= V_t^{\epsilon,1} dt + V_t^{\epsilon,2} dW_t.
 \end{aligned}$$

Now for each $t \in [0, T]$ and each $\epsilon > 0$, we have:

$$\begin{aligned}
 \frac{1}{\epsilon} [b(t, \theta_t^\epsilon) - b(t, \theta_t)] &= \int_0^1 \partial_x b(t, \theta_t^{\lambda, \epsilon}) \cdot (V_t^\epsilon + V_t) d\lambda + \int_0^1 \partial_\alpha b(t, \theta_t^{\lambda, \epsilon}) \cdot \beta_t d\lambda \\
 &+ \int_0^1 \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t^{\lambda, \epsilon})(\tilde{X}_t^{\lambda, \epsilon}) \cdot (\tilde{V}_t^\epsilon + \tilde{V}_t)] d\lambda,
 \end{aligned}$$

where, in order to simplify a little bit the notation, we have set $X_t^{\lambda, \epsilon} = X_t + \lambda\epsilon(V_t^\epsilon + V_t)$, $\alpha_t^{\lambda, \epsilon} = \alpha_t + \lambda\epsilon\beta_t$ and $\theta_t^{\lambda, \epsilon} = (X_t^{\lambda, \epsilon}, \mathbb{P}_{X_t^{\lambda, \epsilon}}, \alpha_t^{\lambda, \epsilon})$. Computing the ‘ dt ’-term, we get:

$$\begin{aligned}
 V_t^{\epsilon,1} &= \int_0^1 \partial_x b(t, \theta_t^{\lambda, \epsilon}) \cdot V_t^\epsilon d\lambda + \int_0^1 \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t^{\lambda, \epsilon})(\tilde{X}_t^{\lambda, \epsilon}) \cdot \tilde{V}_t^\epsilon] d\lambda \\
 &+ \int_0^1 [\partial_x b(t, \theta_t^{\lambda, \epsilon}) - \partial_x b(t, \theta_t)] \cdot V_t d\lambda + \int_0^1 [\partial_\alpha b(t, \theta_t^{\lambda, \epsilon}) - \partial_\alpha b(t, \theta_t)] \cdot \beta_t d\lambda \\
 &+ \int_0^1 \tilde{\mathbb{E}}[(\partial_\mu b(t, \theta_t^{\lambda, \epsilon})(\tilde{X}_t^{\lambda, \epsilon}) - \partial_\mu b(t, \theta_t)(\tilde{X}_t)) \cdot \tilde{V}_t] d\lambda \\
 &= \int_0^1 \partial_x b(t, \theta_t^{\lambda, \epsilon}) \cdot V_t^\epsilon d\lambda + \int_0^1 \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t^{\lambda, \epsilon})(\tilde{X}_t^{\lambda, \epsilon}) \cdot \tilde{V}_t^\epsilon] d\lambda + I_t^{\epsilon,1} + I_t^{\epsilon,2} + I_t^{\epsilon,3}.
 \end{aligned}$$

The three last terms of the above right hand side are bounded in $L^2([0, T] \times \Omega)$, uniformly in ϵ . Indeed, the Lipschitz regularity of b implies that $\partial_x b$ and $\partial_\alpha b$ are bounded and that $\partial_\mu b(t, x, \mathbb{P}_X, \alpha)(X)$ is bounded in $L^2(\Omega; \mathbb{R}^d)$, uniformly in $(t, x, \alpha) \in [0, T] \times \mathbb{R}^d \times A$ and $X \in L^2(\Omega, \mathbb{R}^d)$. Next, we treat the diffusion part $V_t^{\epsilon,2}$ in the same way using Jensen’s inequality and Burkholder-Davis-Gundy’s inequality to control the quadratic variation of the stochastic integrals. Consequently, going back to (27), we see that, for any $S \in [0, T]$,

$$\mathbb{E} \sup_{0 \leq t \leq S} |V_t^\epsilon|^2 \leq c' + c' \int_0^S \mathbb{E} \sup_{0 \leq s \leq t} |V_s^\epsilon|^2 dt,$$

where as usual $c' > 0$ is a generic constant which can change from line to line. Applying Gronwall’s inequality, we deduce that $\mathbb{E} \sup_{0 \leq t \leq T} |V_t^\epsilon|^2 \leq c'$. Therefore, we have

$$\lim_{\epsilon \searrow 0} \mathbb{E} \left[\sup_{0 \leq \lambda \leq 1} \sup_{0 \leq t \leq T} |X_t^{\epsilon, \lambda} - X_t|^2 \right] = 0.$$

We then prove that $I^{1,\epsilon}$, $I^{3,\epsilon}$ and $I^{3,\epsilon}$ converge to 0 in $L^2([0, T] \times \Omega)$ as $\epsilon \searrow 0$. Indeed,

$$\begin{aligned} \mathbb{E} \int_0^T |I_t^{\epsilon,1}|^2 dt &= \mathbb{E} \int_0^T \left| \int_0^1 [\partial_x b(t, \theta_t^{\lambda,\epsilon}) - \partial_x b(t, \theta_t)] V_t d\lambda \right|^2 dt \\ &\leq \mathbb{E} \int_0^T \int_0^1 |\partial_x b(t, \theta_t^{\lambda,\epsilon}) - \partial_x b(t, \theta_t)|^2 |V_t|^2 d\lambda dt. \end{aligned}$$

Since the function $\partial_x b$ is bounded and continuous in x , μ and α , the above right-hand side converges to 0 as $\epsilon \searrow 0$. A similar argument applies to $I_t^{\epsilon,2}$ and $I_t^{\epsilon,3}$. Again, we treat the diffusion part $V_t^{\epsilon,2}$ in the same way using Jensen's inequality and Burkholder-Davis-Gundy's inequality. Consequently, going back to (27), we finally see that, for any $S \in [0, T]$,

$$\mathbb{E} \sup_{0 \leq t \leq S} |V_t^\epsilon|^2 dt \leq c' \int_0^S \mathbb{E} \sup_{0 \leq s \leq t} |V_s^\epsilon|^2 dt + \delta_\epsilon$$

where $\lim_{\epsilon \searrow 0} \delta_\epsilon = 0$. Finally, we get the desired result applying Gronwall's inequality. \square

We now compute the Gâteaux derivative of the objective function.

Lemma 4.2. *The function $\alpha \mapsto J(\alpha)$ is Gâteaux differentiable and its derivative in the direction β is given by:*

$$(28) \quad \begin{aligned} \frac{d}{d\epsilon} J(\alpha + \epsilon\beta) \Big|_{\epsilon=0} &= \mathbb{E} \int_0^T [\partial_x f(t, \theta_t) \cdot V_t + \tilde{\mathbb{E}}[\partial_\mu f(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] + \partial_\alpha f(t, \theta_t) \cdot \beta_t] dt \\ &\quad + \mathbb{E} [\partial_x g(X_T, \mathbb{P}_{X_T}) \cdot V_T + \tilde{\mathbb{E}}[\partial_\mu g(X_T, \mathbb{P}_{X_T})(\tilde{X}_T) \cdot \tilde{V}_T]]. \end{aligned}$$

Proof. We use freely the notation introduced in the proof of the previous lemma.

$$(29) \quad \begin{aligned} \frac{d}{d\epsilon} J(\alpha + \epsilon\beta) \Big|_{\epsilon=0} &= \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \mathbb{E} \int_0^T [f(t, \theta_t^\epsilon) - f(t, \theta_t)] dt \\ &\quad + \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \mathbb{E} [g(X_T^\epsilon, \mathbb{P}_{X_T^\epsilon}) - g(X_T, \mathbb{P}_{X_T})]. \end{aligned}$$

Computing the two limits separately we get

$$\begin{aligned} &\lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \mathbb{E} \int_0^T [f(t, \theta_t^\epsilon) - f(t, \theta_t)] dt \\ &= \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \mathbb{E} \int_0^T \int_0^1 \frac{d}{d\lambda} f(t, \theta_t^{\lambda,\epsilon}) d\lambda dt \\ &= \lim_{\epsilon \searrow 0} \mathbb{E} \int_0^T \int_0^1 [\partial_x f(t, \theta_t^{\lambda,\epsilon}) \cdot (V_t^\epsilon + V_t) \\ &\quad + \tilde{\mathbb{E}}[\partial_\mu f(t, \theta_t^{\lambda,\epsilon})(\tilde{X}_t^{\lambda,\epsilon}) \cdot (\tilde{V}_t^\epsilon + \tilde{V}_t)] + \partial_\alpha f(t, \theta_t^{\lambda,\epsilon}) \cdot \beta_t] d\lambda dt \\ &= \mathbb{E} \int_0^T [\partial_x f(t, \theta_t) \cdot V_t + \tilde{\mathbb{E}}[\partial_\mu f(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] + \partial_\alpha f(t, \theta_t) \cdot \beta_t] dt \end{aligned}$$

using the hypothesis on the continuity and growth of the derivatives of f , the uniform convergence proven in the previous lemma and standard uniform integrability arguments. The second term in (29) is tackled in a similar way. \square

Observing that the conditions (20–21) are satisfied under (A1–4), the duality relationship is given by:

Lemma 4.3. *Given $(Y_t, Z_t)_{0 \leq t \leq T}$ as in Definition 3.5, it holds:*

$$(30) \quad \begin{aligned} \mathbb{E}[Y_T \cdot V_T] &= \mathbb{E} \int_0^T [Y_t \cdot (\partial_\alpha b(t, \theta_t) \cdot \beta_t) + Z_t \cdot (\partial_\alpha \sigma(t, \theta_t) \cdot \beta_t) \\ &\quad - \partial_x f(t, \theta_t) \cdot V_t - \tilde{\mathbb{E}}[\partial_\mu f(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t]] dt. \end{aligned}$$

Proof. Letting $\Theta_t = (X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t)$ and using the definitions (24) of the variation process V , and (22) or (23) of the adjoint process Y , integration by parts gives:

$$\begin{aligned} Y_T \cdot V_T &= Y_0 \cdot V_0 + \int_0^T Y_t \cdot dV_t + \int_0^T dY_t \cdot V_t + \int_0^T d[Y, V]_t \\ &= M_T + \int_0^T \left[Y_t \cdot (\partial_x b(t, \theta_t) \cdot V_t) + Y_t \cdot \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] + Y_t \cdot (\partial_\alpha b(t, \theta_t) \cdot \beta_t) \right. \\ &\quad \left. - \partial_x H(t, \Theta_t) \cdot V_t - \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t) \cdot V_t] \right. \\ &\quad \left. + Z_t \cdot (\partial_x \sigma(t, \theta_t) \cdot V_t) + Z_t \cdot \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] + Z_t \cdot (\partial_\alpha \sigma(t, \theta_t) \cdot \beta_t) \right] dt, \end{aligned}$$

where $(M_t)_{0 \leq t \leq T}$ is a mean zero integrable martingale. By taking expectations on both sides and applying Fubini's theorem:

$$\begin{aligned} &\mathbb{E} \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t) \cdot V_t] \\ &= \mathbb{E} \tilde{\mathbb{E}}[\partial_\mu H(t, \Theta_t)(\tilde{X}_t) \cdot \tilde{V}_t] \\ &= \mathbb{E} \tilde{\mathbb{E}}[(\partial_\mu b(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t) \cdot Y_t + (\partial_\mu \sigma(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t) \cdot Z_t + \partial_\mu f(t, \theta_t)(\tilde{X}_t) \cdot \tilde{V}_t]. \end{aligned}$$

By commutativity of the inner product, cancellations occur and we get the desired equality (30). \square

By putting together the duality relation (30) and (28) we get:

Corollary 4.4. *The Gâteaux derivative of J at α in the direction of β can be written as:*

$$(31) \quad \frac{d}{d\epsilon} J(\alpha + \epsilon\beta)|_{\epsilon=0} = \mathbb{E} \left[\int_0^T \partial_\alpha H(t, X_t, \mathbb{P}_{X_t}, Y_t, \alpha_t) \cdot \beta_t \right] dt.$$

Proof. Using Fubini's theorem, the second expectation appearing in the expression (28) of the Gâteaux derivative of J given in Lemma 4.2 can be rewritten as:

$$\begin{aligned} &\mathbb{E}[\partial_x g(X_T, \mathbb{P}_{X_T}) \cdot V_T + \tilde{\mathbb{E}}(\partial_\mu g(X_T, \mathbb{P}_{X_T})(\tilde{X}_T) \cdot \tilde{V}_T)] \\ &= \mathbb{E}[\partial_x g(X_T, \mathbb{P}_{X_T}) \cdot V_T] + \mathbb{E} \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{\tilde{X}_T})(X_T) \cdot V_T] \\ &= \mathbb{E}[Y_T \cdot V_T], \end{aligned}$$

and using the expression derived in Lemma 4.3 for $\mathbb{E}[Y_T \cdot V_T]$ in (28) gives the desired result. \square

The main result of this subsection is the following theorem.

Theorem 4.5. *Under the above assumptions, if we assume further that the Hamiltonian H is convex in α , the admissible control $\alpha \in \mathbb{A}$ is optimal, X is the associated (optimally) controlled state, and (Y, Z) are the associated adjoint processes solving the adjoint equation (22), then we have:*

$$(32) \quad \forall \alpha \in A, \quad H(t, X_t, Y_t, Z_t, \alpha_t) \leq H(t, X_t, Y_t, Z_t, \alpha) \quad \text{a.e. in } t \in [0, T], \mathbb{P} - \text{a.s.}$$

Proof. Since A is convex, given $\beta \in \mathbb{A}$ we can choose the perturbation $\alpha_t^\epsilon = \alpha_t + \epsilon(\beta_t - \alpha_t)$ which is still in \mathbb{A} for $0 \leq \epsilon \leq 1$. Since α is optimal, we have the inequality

$$\frac{d}{d\epsilon} J(\alpha + \epsilon(\beta - \alpha)) \Big|_{\epsilon=0} = \mathbb{E} \int_0^T \partial_\alpha H(t, X_t, Y_t, Z_t, \alpha_t) \cdot (\beta_t - \alpha_t) \geq 0.$$

By convexity of the Hamiltonian with respect to the control variable $\alpha \in A$, we conclude that

$$\mathbb{E} \int_0^T [H(t, X_t, Y_t, Z_t, \beta_t) - H(t, X_t, Y_t, Z_t, \alpha_t)] dt \geq 0,$$

for all β . Now, if for a given (deterministic) $\alpha \in A$ we choose β in the following way,

$$\beta_t(\omega) = \begin{cases} \alpha & \text{if } (t, \omega) \in C \\ \alpha_t(\omega) & \text{otherwise} \end{cases}$$

for an arbitrary progressively-measurable set $C \subset [0, T] \times \Omega$ (that is $C \cap [0, t] \in \mathcal{B}([0, t]) \otimes \mathcal{F}_t$ for any $t \in [0, T]$), we see that

$$\mathbb{E} \int_0^T \mathbf{1}_C [H(t, X_t, Y_t, Z_t, \alpha) - H(t, X_t, Y_t, Z_t, \alpha_t)] dt \geq 0,$$

from which we conclude that

$$H(t, X_t, Y_t, Z_t, \alpha) - H(t, X_t, Y_t, Z_t, \alpha_t) \geq 0 \quad dt \otimes d\mathbb{P} \text{ a.e.},$$

which is the desired conclusion. \square

4.2. A Sufficient Condition. The necessary condition for optimality identified in the previous subsection can be turned into a sufficient condition for optimality under some technical assumptions.

Theorem 4.6. *Under the same assumptions of regularity on the coefficients as before, let $\alpha \in \mathbb{A}$ be an admissible control, $X = X^\alpha$ the corresponding controlled state process, and (Y, Z) the corresponding adjoint processes. Let us also assume that for each $t \in [0, T]$*

- (1) $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, \mu) \mapsto g(x, \mu)$ is convex;
- (2) $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \ni (x, \mu, \alpha) \mapsto H(t, x, \mu, Y_t, Z_t, \alpha)$ is convex $dt \otimes d\mathbb{P}$ almost everywhere.

Moreover, if

$$(33) \quad H(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t) = \inf_{\alpha \in A} H(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha), \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

then α is an optimal control, i.e. $J(\alpha) = \inf_{\alpha' \in \mathbb{A}} J(\alpha')$.

Proof. Let $\alpha' \in \mathbb{A}$ be a generic admissible control, and $X' = X^{\alpha'}$ the corresponding controlled state. By definition of the objective function of the control problem we have:

$$\begin{aligned}
 J(\alpha) - J(\alpha') &= \mathbb{E}[g(X_T, \mathbb{P}_{X_T}) - g(X'_T, \mathbb{P}_{X'_T})] + \mathbb{E} \int_0^T [f(t, \theta_t) - f(t, \theta'_t)] dt \\
 (34) \quad &= \mathbb{E}[g(X_T, \mathbb{P}_{X_T}) - g(X'_T, \mathbb{P}_{X'_T})] + \mathbb{E} \int_0^T [H(t, \Theta_t) - H(t, \Theta'_t)] dt \\
 &\quad - \mathbb{E} \int_0^T \{ [b(t, \theta_t) - b(t, \theta'_t)] \cdot Y_t + [\sigma(t, \theta_t) - \sigma(t, \theta'_t)] \cdot Z_t \} dt
 \end{aligned}$$

by definition of the Hamiltonian, where $\theta_t = (X_t, \mathbb{P}_{X_t}, \alpha_t)$ and $\Theta_t = (X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t)$ (and similarly for θ'_t and Θ'_t). The function g being convex, we have

$$g(x, \mu) - g(x', \mu') \leq (x - x') \cdot \partial_x g(x) + \tilde{\mathbb{E}}[\partial_\mu g(x, \mu)(\tilde{X}) \cdot (\tilde{X} - \tilde{X}')],$$

so that

$$\begin{aligned}
 &\mathbb{E}[g(X_T, \mathbb{P}_{X_T}) - g(X'_T, \mathbb{P}_{X'_T})] \\
 (35) \quad &\leq \mathbb{E}[\partial_x g(X_T, \mathbb{P}_{X_T}) \cdot (X_T - X'_T) + \tilde{\mathbb{E}}[\partial_\mu g(X_T, \mathbb{P}_{X_T})(\tilde{X}_T) \cdot (\tilde{X}_T - \tilde{X}'_T)]] \\
 &= \mathbb{E}[(\partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]) \cdot (X_T - X'_T)] \\
 &= \mathbb{E}[Y_T \cdot (X_T - X'_T)] = \mathbb{E}[(X_T - X'_T) \cdot Y_T],
 \end{aligned}$$

where we used Fubini's theorem and the fact that the 'tilde random variables' are independent copies of the 'non-tilde variables'. Using the adjoint equation and taking the expectation, we get:

$$\begin{aligned}
 &\mathbb{E}[(X_T - X'_T) \cdot Y_T] \\
 &= \mathbb{E} \left[\int_0^T (X_t - X'_t) \cdot dY_t + \int_0^T Y_t \cdot d[X_t - X'_t] + \int_0^T [\sigma(t, \theta_t) - \sigma(t, \theta'_t)] \cdot Z_t dt \right] \\
 &= -\mathbb{E} \int_0^T [\partial_x H(t, \Theta_t) \cdot (X_t - X'_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t)] \cdot (X_t - X'_t)] dt \\
 &\quad + \mathbb{E} \int_0^T [b(t, \theta_t) - b(t, \theta'_t)] \cdot Y_t + [\sigma(t, \theta_t) - \sigma(t, \theta'_t)] \cdot Z_t dt,
 \end{aligned}$$

where we used integration by parts and the fact that Y_t solves the adjoint equation. Using Fubini's theorem and the fact that $\tilde{\Theta}_t$ is an independent copy of Θ_t , the expectation of the second term in the second line can be rewritten as

$$\begin{aligned}
 (36) \quad &\mathbb{E} \int_0^T \{ \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t)] \cdot (X_t - X'_t) \} dt = \mathbb{E} \tilde{\mathbb{E}} \int_0^T \{ [\partial_\mu H(t, \Theta_t)(\tilde{X}_t)] \cdot (\tilde{X}_t - \tilde{X}'_t) \} dt \\
 &= \mathbb{E} \int_0^T \tilde{\mathbb{E}}[\partial_\mu H(t, \Theta_t)(\tilde{X}_t) \cdot (\tilde{X}_t - \tilde{X}'_t)] dt.
 \end{aligned}$$

Consequently, by (34), (35) and (36), we obtain

$$\begin{aligned}
 J(\alpha) - J(\alpha') &\leq \mathbb{E} \int_0^T [H(t, \Theta_t) - H(t, \Theta'_t)] dt \\
 (37) \quad &- \mathbb{E} \int_0^T \{ \partial_x H(t, \Theta_t) \cdot (X_t - X'_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t) \cdot (\tilde{X}_t - \tilde{X}'_t)] \} dt \\
 &\leq 0
 \end{aligned}$$

because of the convexity assumption on H , see in particular (13), and because of the criticality of the admissible control $(\alpha_t)_{0 \leq t \leq T}$, see (33), which says the first order derivative in α vanishes. \square

4.3. Special Cases. We consider a set of particular cases which already appeared in the literature, and we provide the special forms of the Pontryagin maximum principle which apply in these cases. We discuss only sufficient conditions for optimality for the sake of definiteness. The corresponding necessary conditions can easily be derived from the results of Subsection 4.1.

Scalar Interactions. In this subsection we show how the model handled in [2] appears as a specific example of our more general formulation. We consider scalar interactions for which the dependence upon the probability measure of the coefficients of the dynamics of the state and the cost functions is through functions of scalar moments of the measure. More specifically, we assume that:

$$\begin{aligned}
 b(t, x, \mu, \alpha) &= \hat{b}(t, x, \langle \psi, \mu \rangle, \alpha) & \sigma(t, x, \mu, \alpha) &= \hat{\sigma}(t, x, \langle \phi, \mu \rangle, \alpha) \\
 f(t, x, \mu, \alpha) &= \hat{f}(t, x, \langle \gamma, \mu \rangle, \alpha) & g(x, \mu) &= \hat{g}(x, \langle \zeta, \mu \rangle)
 \end{aligned}$$

for some scalar functions ψ, ϕ, γ and ζ with at most quadratic growth at ∞ , and functions $\hat{b}, \hat{\sigma}$ and \hat{f} defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R} \times A$ with values in $\mathbb{R}^d, \mathbb{R}^{d \times m}$ and \mathbb{R} respectively, and a real valued function \hat{g} defined on $\mathbb{R}^d \times \mathbb{R}$. We use the bracket notation $\langle h, \mu \rangle$ to denote the integral of the function h with respect to the measure μ . The functions $\hat{b}, \hat{\sigma}, \hat{f}$ and \hat{g} are similar to the functions b, σ, f and g with the variable μ , which was a measure, replaced by a numeric variable, say r . Reserving the notation H for the Hamiltonian we defined above, we have:

$$H(t, x, \mu, y, z, \alpha) = \hat{b}(t, x, \langle \psi, \mu \rangle, \alpha) \cdot y + \hat{\sigma}(t, x, \langle \phi, \mu \rangle, \alpha) \cdot z + \hat{f}(t, x, \langle \gamma, \mu \rangle, \alpha).$$

We then proceed to derive the particular form taken by the adjoint equation in the present situation. We start with the terminal condition as it is easier to identify. According to (22), it reads:

$$Y_T = \partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{\tilde{X}_T})(X_T)].$$

Since the terminal cost is of the form $g(x, \mu) = \hat{g}(x, \langle \zeta, \mu \rangle)$, given our definition of differentiability with respect to the variable μ , we know, as a generalization of (6), that $\partial_\mu g(x, \mu)(\cdot)$ reads

$$\partial_\mu g(x, \mu)(x') = \partial_r \hat{g}(x, \langle \zeta, \mu \rangle) \partial \zeta(x'), \quad x' \in \mathbb{R}^d.$$

Therefore, the terminal condition Y_T can be rewritten as

$$Y_T = \partial_x \hat{g}(X_T, \mathbb{E}[\zeta(X_T)]) + \tilde{\mathbb{E}}[\partial_r \hat{g}(\tilde{X}_T, \mathbb{E}[\zeta(X_T)])] \partial \zeta(X_T)$$

which is exactly the terminal condition used in [2] once we remark that the ‘tildes’ can be removed since \tilde{X}_T has the same distribution as X_T . Within this framework, convexity in μ is quite easy to check. Here is a typical example borrowed from [2]: if g and \hat{g} do not depend on x , then the function $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto g(\mu) = \hat{g}(\langle \zeta, \mu \rangle)$ is convex if ζ is convex and \hat{g} is non-decreasing and convex.

Similarly, $\partial_\mu H(t, x, \mu, y, z, \alpha)$ can be identified to the \mathbb{R}^d -valued function defined by

$$\begin{aligned} \partial_\mu H(t, x, \mu, y, z, \alpha)(x') &= [\partial_r \hat{b}(t, x, \langle \psi, \mu \rangle, \alpha) \odot y] \partial \psi(x') + [\partial_r \hat{\sigma}(t, x, \langle \phi, \mu \rangle, \alpha) \odot z] \partial \phi(x') \\ &\quad + \partial_r \hat{f}(t, x, \langle \gamma, \mu \rangle, \alpha) \partial \gamma(x') \end{aligned}$$

and the dynamic part of the adjoint equation (22) rewrites:

$$\begin{aligned} dY_t &= -\left\{ \partial_x \hat{b}(t, X_t, \mathbb{E}[\psi(X_t)], \alpha_t) \odot Y_t + \partial_x \hat{\sigma}(t, X_t, \mathbb{E}[\phi(X_t)], \alpha_t) \odot Z_t \right. \\ &\quad \left. + \partial_x \hat{f}(t, X_t, \mathbb{E}[\gamma(X_t)], \alpha_t) \right\} dt + Z_t dW_t \\ &\quad - \left\{ \tilde{\mathbb{E}}[\partial_r \hat{b}(t, \tilde{X}_t, \mathbb{E}[\psi(X_t)], \tilde{\alpha}_t) \odot \tilde{Y}_t] \partial \psi(X_t) + \tilde{\mathbb{E}}[\partial_r \hat{\sigma}(t, \tilde{X}_t, \mathbb{E}[\phi(X_t)], \tilde{\alpha}_t) \odot \tilde{Z}_t] \partial \phi(X_t) \right. \\ &\quad \left. + \tilde{\mathbb{E}}[\partial_r \hat{f}(t, \tilde{X}_t, \mathbb{E}[\gamma(X_t)], \tilde{\alpha}_t)] \partial \gamma(X_t) \right\} dt, \end{aligned}$$

which again, is exactly the adjoint equation used in [2] once we *remove the 'tildes'*.

Remark 4.7. *The mean variance portfolio optimization example discussed in [2] and the solution proposed in [3] and [8] of the optimal control of linear-quadratic (LQ) McKean-Vlasov dynamics are based on the general form of the Pontryagin principle proven in this section as applied to the scalar interactions considered in this subsection.*

First Order Interactions. In the case of first order interactions, the dependence upon the probability measure is linear in the sense that the coefficients b , σ , f and g are given in the form

$$\begin{aligned} b(t, x, \mu, \alpha) &= \langle \hat{b}(t, x, \cdot, \alpha), \mu \rangle & \sigma(t, x, \mu, \alpha) &= \langle \hat{\sigma}(t, x, \cdot, \alpha), \mu \rangle \\ f(t, x, \mu, \alpha) &= \langle \hat{f}(t, x, \cdot, \alpha), \mu \rangle & g(x, \mu) &= \langle \hat{g}(x, \cdot), \mu \rangle \end{aligned}$$

for some functions \hat{b} , $\hat{\sigma}$ and \hat{f} defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times A$ with values in \mathbb{R}^d , $\mathbb{R}^{d \times m}$ and \mathbb{R} respectively, and a real valued function \hat{g} defined on $\mathbb{R}^d \times \mathbb{R}^d$. The form of this dependence comes from the original derivation of the McKean-Vlasov equation as limit of the dynamics of a large system of particles evolving according to a system of stochastic differential equations with *mean field* interactions of the form

$$(38) \quad dX_t^i = \frac{1}{N} \sum_{j=1}^N \hat{b}(t, X_t^i, X_t^j) dt + \frac{1}{N} \sum_{j=1}^N \hat{\sigma}(t, X_t^i, X_t^j) dW_t^j, \quad i = 1, \dots, N, \quad 0 \leq t \leq T,$$

where W^i 's are N independent standard Wiener processes in \mathbb{R}^d . In the present situation the linearity in μ implies that $\partial_\mu g(x, \mu)(x') = \partial_{x'} \hat{g}(x, x')$ and similarly

$$\partial_\mu H(t, x, \mu, y, z, \alpha)(x') = \partial_{x'} \hat{b}(t, x, x', \alpha) \odot y + \partial_{x'} \hat{\sigma}(t, x, x', \alpha) \odot z + \partial_{x'} \hat{f}(t, x, x', \alpha),$$

and the dynamic part of the adjoint equation (22) rewrites:

$$dY_t = -\tilde{\mathbb{E}}[\partial_x \hat{H}(t, X_t, \tilde{X}_t, Y_t, Z_t, \alpha_t) + \partial_{x'} \hat{H}(t, \tilde{X}_t, X_t, \tilde{Y}_t, \tilde{Z}_t, \tilde{\alpha}_t)] dt + Z_t dW_t,$$

if we use the obvious notation

$$\hat{H}(t, x, x', y, z, \alpha) = \hat{b}(t, x, x', \alpha) \cdot y + \hat{\sigma}(t, x, x', \alpha) \cdot z + \hat{f}(t, x, x', \alpha),$$

and the terminal condition is given by

$$Y_T = \tilde{\mathbb{E}}[\partial_x \hat{g}(X_T, \tilde{X}_T) + \partial_{x'} \hat{g}(\tilde{X}_T, X_T)].$$

5. SOLVABILITY OF FORWARD-BACKWARD SYSTEMS

We now turn to the application of the stochastic maximum principle to the solution of the optimal control of McKean-Vlasov dynamics. The strategy is to identify a minimizer of the Hamiltonian, and to use it in the forward dynamics and the adjoint equation. This creates a coupling between these equations, leading to the study of an FBSDE of mean field type. As explained in the introduction, the existence results proven in [7] and [6] do not cover some of the solvable models (such as the LQ models). Here we establish existence and uniqueness by taking advantage of the specific structure of the equation, inherited from the underlying optimization problem. Assuming that the terminal cost and the Hamiltonian satisfy the same convexity assumptions as in the statement of Theorem 4.6, we indeed prove that unique solvability holds by applying the *continuation method*, originally exposed within the framework of FBSDEs in [14]. Some of the results of this section were announced in the note [8].

5.1. Technical Assumptions. We now formulate the assumptions we shall use from now on. These assumptions subsume the assumptions (A1-4) introduced in Sections 2 and 4. As it is most often the case in applications of Pontryagin's stochastic maximum principle, we choose $A = \mathbb{R}^k$ and we consider a *linear* model for the forward dynamics of the state.

(B1) The drift b and the volatility σ are linear in μ , x and α . They read

$$\begin{aligned} b(t, x, \mu, \alpha) &= b_0(t) + b_1(t)\bar{\mu} + b_2(t)x + b_3(t)\alpha, \\ \sigma(t, x, \mu, \alpha) &= \sigma_0(t) + \sigma_1(t)\bar{\mu} + \sigma_2(t)x + \sigma_3(t)\alpha, \end{aligned}$$

for some bounded measurable deterministic functions b_0, b_1, b_2 and b_3 with values in $\mathbb{R}^d, \mathbb{R}^{d \times d}, \mathbb{R}^{d \times d}$ and $\mathbb{R}^{d \times k}$ and $\sigma_0, \sigma_1, \sigma_2$ and σ_3 with values in $\mathbb{R}^{d \times m}, \mathbb{R}^{(d \times m) \times d}, \mathbb{R}^{(d \times m) \times d}$ and $\mathbb{R}^{(d \times m) \times k}$ (the parentheses around $d \times m$ indicating that $\sigma_i(t)u_i$ is seen as an element of $\mathbb{R}^{d \times m}$ whenever $u_i \in \mathbb{R}^d$, with $i = 1, 2$, or $u_i \in \mathbb{R}^k$, with $i = 3$), and where we use the notation $\bar{\mu} = \int x d\mu(x)$ for the mean of a measure μ .

(B2) The functions f and g satisfy the same assumptions as in (A.3-4) in Section 4 (with respect to some constant L). In particular, there exists a constant \hat{L} such that

$$\begin{aligned} &|f(t, x', \mu', \alpha') - f(t, x, \mu, \alpha)| + |g(x', \mu') - g(x, \mu)| \\ &\leq \hat{L}[1 + |x'| + |x| + |\alpha'| + |\alpha| + \|\mu\|_2 + \|\mu'\|_2][|(x', \alpha') - (x, \alpha)| + W_2(\mu', \mu)]. \end{aligned}$$

(B3) There exists a constant $\hat{c} > 0$ such that the derivatives of f and g with respect to (x, α) and x respectively are \hat{c} -Lipschitz continuous with respect to (x, α, μ) and (x, μ) respectively (the Lipschitz property in the variable μ being understood in the sense of the 2-Wasserstein distance). Moreover, for any $t \in [0, T]$, any $x, x' \in \mathbb{R}^d$, any $\alpha, \alpha' \in \mathbb{R}^k$, any $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and any \mathbb{R}^d -valued random variables X and X' having μ and μ' as respective distributions,

$$\begin{aligned} \mathbb{E}[|\partial_\mu f(t, x', \mu', \alpha')(X') - \partial_\mu f(t, x, \mu, \alpha)(X)|^2] &\leq \hat{c}(|(x', \alpha') - (x, \alpha)|^2 + \mathbb{E}[|X' - X|^2]), \\ \mathbb{E}[|\partial_\mu g(x', \mu')(X') - \partial_\mu g(x, \mu)(X)|^2] &\leq \hat{c}(|x' - x|^2 + \mathbb{E}[|X' - X|^2]). \end{aligned}$$

(B4) The function f is convex with respect to (x, μ, α) for t fixed in such a way that, for some $\lambda > 0$,

$$\begin{aligned} & f(t, x', \mu', \alpha') - f(t, x, \mu, \alpha) \\ & - \partial_{(x, \alpha)} f(t, x, \mu, \alpha) \cdot (x' - x, \alpha' - \alpha) - \tilde{\mathbb{E}}[\partial_\mu f(t, x, \mu, \alpha)(\tilde{X}) \cdot (\tilde{X}' - \tilde{X})] \\ & \geq \lambda |\alpha' - \alpha|^2, \end{aligned}$$

whenever $\tilde{X}, \tilde{X}' \in L^2(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}}; \mathbb{R}^d)$ with distributions μ and μ' respectively. The function g is also assumed to be convex in (x, μ) (on the same model, but with $\lambda = 0$).

We refer to Subsection 3.2 for a precise discussion about (B3) and (B4). By comparing (7) with (B3), notice in particular that the liftings $L^2(\tilde{\Omega}; \mathbb{R}^d) \ni \tilde{X} \mapsto f(t, x, \tilde{\mathbb{P}}_{\tilde{X}}, \alpha)$ and $L^2(\tilde{\Omega}; \mathbb{R}^d) \ni \tilde{X} \mapsto g(x, \tilde{\mathbb{P}}_{\tilde{X}})$ have Lipschitz continuous derivatives. As a consequence, Lemma 3.2 applies: For any $t \in [0, T]$, $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\alpha \in \mathbb{R}^k$, there exist versions of $\mathbb{R}^d \ni x' \mapsto \partial_\mu f(t, x, \mu, \alpha)(x')$ and $\mathbb{R}^d \ni x' \mapsto \partial_\mu g(x, \mu)(x')$ that are \hat{c} -Lipschitz continuous.

Following Example (6), we also emphasize that b and σ obviously satisfy (B3).

5.2. The Hamiltonian and the Adjoint Equations. The drift and the volatility being linear, the Hamiltonian has the form

$$\begin{aligned} H(t, x, \mu, y, z, \alpha) &= [b_0(t) + b_1(t)\bar{\mu} + b_2(t)x + b_3(t)\alpha] \cdot y \\ &+ [\sigma_0(t) + \sigma_1(t)\bar{\mu} + \sigma_2(t)x + \sigma_3(t)\alpha] \cdot z + f(t, x, \mu, \alpha), \end{aligned}$$

for $t \in [0, T]$, $x, y \in \mathbb{R}^d$, $z \in \mathbb{R}^{d \times m}$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\alpha \in \mathbb{R}^k$. Given $(t, x, \mu, y, z) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d \times m}$, the function $\mathbb{R}^k \ni \alpha \mapsto H(t, x, \mu, y, z, \alpha)$ is strictly convex so that there exists a unique minimizer $\hat{\alpha}(t, x, \mu, y, z)$:

$$(39) \quad \hat{\alpha}(t, x, \mu, y, z) = \operatorname{argmin}_{\alpha \in \mathbb{R}^k} H(t, x, \mu, y, z, \alpha).$$

Assumptions (B1-4) above being slightly stronger than the assumptions used in [7], we can follow the arguments given in the proof of Lemma 2.1 of [7] in order to prove that, for all $(t, x, \mu, y, z) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d \times m}$, the function $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d \times m} \ni (t, x, \mu, y, z) \mapsto \hat{\alpha}(t, x, \mu, y, z)$ is measurable, locally bounded and Lipschitz-continuous with respect to (x, μ, y, z) , uniformly in $t \in [0, T]$, the Lipschitz constant depending only upon λ , the supremum norms of b_3 and σ_3 and the Lipschitz constant of $\partial_\alpha f$ in (x, μ) . Except maybe for the Lipschitz property with respect to the measure argument, these facts were explicitly proved in [7]. The regularity of $\hat{\alpha}$ with respect to μ follows from the following remark. If $(t, x, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d \times m}$ is fixed and μ, μ' are generic elements in $\mathcal{P}_2(\mathbb{R}^d)$, $\hat{\alpha}$ and $\hat{\alpha}'$ denoting the associated minimizers, we deduce from the convexity assumption (B4):

$$\begin{aligned} 2\lambda |\hat{\alpha}' - \hat{\alpha}|^2 &\leq (\hat{\alpha}' - \hat{\alpha}) \cdot [\partial_\alpha f(t, x, \mu, \hat{\alpha}') - \partial_\alpha f(t, x, \mu, \hat{\alpha})] \\ &= (\hat{\alpha}' - \hat{\alpha}) \cdot [\partial_\alpha H(t, x, \mu, y, z, \hat{\alpha}') - \partial_\alpha H(t, x, \mu, y, z, \hat{\alpha})] \\ (40) \quad &= (\hat{\alpha}' - \hat{\alpha}) \cdot [\partial_\alpha H(t, x, \mu, y, z, \hat{\alpha}') - \partial_\alpha H(t, x, \mu', y, z, \hat{\alpha}')] \\ &= (\hat{\alpha}' - \hat{\alpha}) \cdot [\partial_\alpha f(t, x, \mu, \hat{\alpha}') - \partial_\alpha f(t, x, \mu', \hat{\alpha}')] \\ &\leq C |\hat{\alpha}' - \hat{\alpha}| W_2(\mu', \mu), \end{aligned}$$

the passage from the second to the third following from the identity

$$\partial_\alpha H(t, x, \mu, y, z, \hat{\alpha}) = \partial_\alpha H(t, x, \mu', y, z, \hat{\alpha}') = 0.$$

For each admissible control $\alpha = (\alpha_t)_{0 \leq t \leq T}$, if we denote the corresponding solution of the state equation by $X = (X_t^\alpha)_{0 \leq t \leq T}$, then the adjoint BSDE (22) introduced in Definition 3.5 reads:

$$(41) \quad \begin{aligned} dY_t = & -\partial_x f(t, X_t, \mathbb{P}_{X_t}, \alpha_t) dt - b_2^\dagger(t) Y_t dt - \sigma_2^\dagger(t) Z_t dt + Z_t dW_t \\ & - \tilde{\mathbb{E}}[\partial_\mu f(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{\alpha}_t)(X_t)] dt - b_1^\dagger(t) \mathbb{E}[Y_t] - \sigma_1^\dagger(t) \mathbb{E}[Z_t] dt. \end{aligned}$$

Given the necessary and sufficient conditions proven in the previous section, our goal is to use the control $(\hat{\alpha}_t)_{0 \leq t \leq T}$ defined by $\hat{\alpha}_t = \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t)$ where $\hat{\alpha}$ is the minimizer function constructed above and the process $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$ is a solution of the FBSDE

$$(42) \quad \begin{cases} dX_t = [b_0(t) + b_1(t) \mathbb{E}[X_t] + b_2(t) X_t + b_3(t) \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t)] dt \\ \quad + [\sigma_0(t) + \sigma_1(t) \mathbb{E}[X_t] + \sigma_2(t) X_t + \sigma_3(t) \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t)] dW_t, \\ dY_t = -[\partial_x f(t, X_t, \mathbb{P}_{X_t}, \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t)) + b_2^\dagger(t) Y_t + \sigma_2^\dagger(t) Z_t] dt + Z_t dW_t \\ \quad - \{\tilde{\mathbb{E}}[\partial_\mu f(t, \tilde{X}_t, \mathbb{P}_{X_t}, \hat{\alpha}(t, \tilde{X}_t, \mathbb{P}_{X_t}, \tilde{Y}_t, \tilde{Z}_t))(X_t)] + b_1^\dagger(t) \mathbb{E}[Y_t] + \sigma_1^\dagger(t) \mathbb{E}[Z_t]\} dt, \end{cases}$$

with the initial condition $X_0 = x_0$, for a given deterministic point $x_0 \in \mathbb{R}^d$, and the terminal condition $Y_T = \partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]$.

5.3. Main Result. Here is the main existence and uniqueness result:

Theorem 5.1. *Under (B1-4), the forward-backward system (42) is uniquely solvable.*

Proof. The proof is an adaptation of the *continuation method* used in [14] to handle standard FBSDEs satisfying appropriate monotonicity conditions. Generally speaking, it consists in proving that existence and uniqueness are kept preserved when the coefficients in (42) are slightly perturbed. Starting from an initial case for which existence and uniqueness are known to hold, we then establish Theorem 5.1 by modifying iteratively the coefficients so that (42) is eventually shown to belong to the class of uniquely solvable systems.

A natural and simple strategy then consists in modifying the coefficients in a linear way. Unfortunately, this might generate heavy notations. For that reason, we use the following conventions.

First, as in Subsection 4.1, the notation $(\Theta_t)_{0 \leq t \leq T}$ stands for the generic notation for denoting a process of the form $(X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t)_{0 \leq t \leq T}$ with values in $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d \times m} \times \mathbb{R}^k$. We will then denote by \mathbb{S} the space of processes $(\Theta_t)_{0 \leq t \leq T}$ such that $(X_t, Y_t, Z_t, \alpha_t)_{0 \leq t \leq T}$ is $(\mathcal{F}_t)_{0 \leq t \leq T}$ progressively-measurable, $(X_t)_{0 \leq t \leq T}$ and $(Y_t)_{0 \leq t \leq T}$ have continuous trajectories, and

$$(43) \quad \|\Theta\|_{\mathbb{S}} = \mathbb{E} \left[\sup_{0 \leq t \leq T} [|X_t|^2 + |Y_t|^2] + \int_0^T [|Z_t|^2 + |\alpha_t|^2] dt \right]^{1/2} < +\infty.$$

Similarly, the notation $(\theta_t)_{0 \leq t \leq T}$ is the generic notation for denoting a process $(X_t, \mathbb{P}_{X_t}, \alpha_t)_{0 \leq t \leq T}$ with values in $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^k$. All the processes $(\theta_t)_{0 \leq t \leq T}$ that are considered below appear as the restrictions of an *extended* process $(\Theta_t)_{0 \leq t \leq T} \in \mathbb{S}$.

Moreover, we call an initial condition for (42) a square-integrable \mathcal{F}_0 -measurable random variable ξ with values in \mathbb{R}^d , that is an element of $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$. Recall indeed that \mathcal{F}_0 can be chosen as a

σ -algebra independent of $(W_t)_{0 \leq t \leq T}$. In comparison with the statement of Theorem 5.1, this permits to generalize the case when ξ is deterministic.

Finally, we call an input for (42) a four-tuple $\mathcal{I} = ((\mathcal{I}_t^b, \mathcal{I}_t^\sigma, \mathcal{I}_t^f)_{0 \leq t \leq T}, \mathcal{I}_T^g)$, $(\mathcal{I}^b)_{0 \leq t \leq T}$, $(\mathcal{I}^\sigma)_{0 \leq t \leq T}$ and $(\mathcal{I}^f)_{0 \leq t \leq T}$ being three square integrable progressively-measurable processes with values in \mathbb{R}^d , $\mathbb{R}^{d \times m}$ and \mathbb{R}^d respectively, and \mathcal{I}^g denoting a square-integrable \mathcal{F}_T -measurable random variable with values in \mathbb{R}^d . Such an input is specifically designed to be injected into the dynamics of (42), \mathcal{I}^b being plugged into the drift of the forward equation, \mathcal{I}^σ into the volatility of the forward equation, \mathcal{I}^f into the bounded variation term of the backward equation and \mathcal{I}^g into the terminal condition of the backward equation. The space of inputs is denoted by \mathbb{I} . It is endowed with the norm:

$$(44) \quad \|\mathcal{I}\|_{\mathbb{I}} = \mathbb{E} \left[|\mathcal{I}_T^g|^2 + \int_0^T [|\mathcal{I}_t^b|^2 + |\mathcal{I}_t^\sigma|^2 + |\mathcal{I}_t^f|^2] dt \right]^{1/2}.$$

We then put:

Definition 5.2. For any $\gamma \in [0, 1]$, any $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ and any input $\mathcal{I} \in \mathbb{I}$, the FBSDE

$$(45) \quad \begin{aligned} dX_t &= (\gamma b(t, \theta_t) + \mathcal{I}_t^b) dt + (\gamma \sigma(t, \theta_t) + \mathcal{I}_t^\sigma) dW_t, \\ dY_t &= -(\gamma \{\partial_x H(t, \Theta_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t)]\} + \mathcal{I}_t^f) dt + Z_t dW_t, \quad t \in [0, T], \end{aligned}$$

with the optimality condition

$$(46) \quad \alpha_t = \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t), \quad t \in [0, T],$$

and with $X_0 = \xi$ as initial condition and

$$Y_T = \gamma \{\partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]\} + \mathcal{I}_T^g$$

as terminal condition, is referred to as $\mathcal{E}(\gamma, \xi, \mathcal{I})$.

Whenever $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$ is a solution, the full process $(X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t)_{0 \leq t \leq T}$ is referred to as the associated extended solution.

Remark 5.3. The way the coupling is summarized between the forward and backward equations in (45) is a bit different from the way Equation (42) is written. In the formulation used in the statement of Lemma 5.2, the coupling between the forward and the backward equations follows from the optimality condition (46). Because of that optimality condition, the two formulations are equivalent: When $\gamma = 1$ and $\mathcal{I} \equiv 0$, the pair (45–46) coincides with (42).

The following lemma is proved in the next subsection:

Lemma 5.4. Given $\gamma \in [0, 1]$, we say the property (\mathcal{S}_γ) holds true if, for any $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ and any $\mathcal{I} \in \mathbb{I}$, the FBSDE $\mathcal{E}(\gamma, \xi, \mathcal{I})$ has a unique extended solution in \mathbb{S} . With this definition, there exists $\delta_0 > 0$ such that, if (\mathcal{S}_γ) holds true for some $\gamma \in [0, 1)$, then $(\mathcal{S}_{\gamma+\eta})$ holds true for any $\eta \in (0, \delta_0]$ satisfying $\gamma + \eta \leq 1$.

Given Lemma 5.4, Theorem 5.1 follows from a straightforward induction as (\mathcal{S}_0) obviously holds true. \square

5.4. Proof of Lemma 5.4. The proof follows from Picard's contraction theorem. As in the statement, consider indeed γ such that (\mathcal{S}_γ) holds true. For $\eta > 0$, $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ and $\mathcal{I} \in \mathbb{I}$, we then define a mapping Φ from \mathbb{S} into itself whose fixed points coincide with the solutions of $\mathcal{E}(\gamma + \eta, \xi, \mathcal{I})$.

The definition of Φ is as follows. Given a process $\Theta \in \mathbb{S}$, we denote by Θ' the extended solution of the FBSDE $\mathcal{E}(\gamma, \xi, \mathcal{I}')$ with

$$\begin{aligned}\mathcal{I}_t^{b,\prime} &= \eta b(t, \theta_t) + \mathcal{I}_t^b, \\ \mathcal{I}_t^{\sigma,\prime} &= \eta \sigma(t, \theta_t) + \mathcal{I}_t^\sigma, \\ \mathcal{I}_t^{f,\prime} &= \eta \partial_x H(t, \Theta_t) + \eta \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t)] + \mathcal{I}_t^f, \\ \mathcal{I}_T^{g,\prime} &= \eta \partial_x g(X_T, \mathbb{P}_{X_T}) + \eta \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)] + \mathcal{I}_T^g.\end{aligned}$$

By assumption, it is uniquely defined and it belongs to \mathbb{S} , so that the mapping $\Phi : \Theta \mapsto \Theta'$ maps \mathbb{S} into itself. It is then clear that a process $\Theta \in \mathbb{S}$ is a fixed point of Φ if and only if Θ is an extended solution of $\mathcal{E}(\gamma + \eta, \xi, \mathcal{I})$. The point is thus to prove that Φ is a contraction when η is small enough. This is a consequence of the following lemma:

Lemma 5.5. *Let $\gamma \in [0, 1]$ such that (\mathcal{S}_γ) holds true. Then, there exists a constant C , independent of γ , such that, for any $\xi, \xi' \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ and $\mathcal{I}, \mathcal{I}' \in \mathbb{I}$, the respective extended solutions Θ and Θ' of $\mathcal{E}(\gamma, \xi, \mathcal{I})$ and $\mathcal{E}(\gamma, \xi', \mathcal{I}')$ satisfy:*

$$\|\Theta - \Theta'\|_{\mathbb{S}} \leq C(\mathbb{E}[|\xi - \xi'|^2]^{1/2} + \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{I}}).$$

Given Lemma 5.5, we indeed check that Φ is a contraction when η is small enough. Given Θ^1 and Θ^2 two processes in \mathbb{S} and denoting by Θ'^1 and Θ'^2 their respective images by Φ , we deduce from Lemma 5.5 that

$$\|\Theta'^1 - \Theta'^2\|_{\mathbb{S}} \leq C\eta \|\Theta^1 - \Theta^2\|_{\mathbb{S}},$$

which is enough to conclude.

5.5. Proof of Lemma 5.5. The strategy follows from a mere variation on the proof of the classical stochastic maximum principle. With the same notations as in the statement and with the classical convention for expanding Θ as $(X_t, \mathbb{P}_{X_t}, Y_t, Z_t, \alpha_t)_{0 \leq t \leq T}$ and for letting $(\theta_t = (X_t, \mathbb{P}_{X_t}, \alpha_t))_{0 \leq t \leq T}$, we then compute

$$\begin{aligned}\mathbb{E}[(X'_T - X_T) \cdot Y_T] &= \mathbb{E}[(\xi' - \xi) \cdot Y_0] \\ &\quad - \gamma \left\{ \mathbb{E} \int_0^T [\partial_x H(t, \Theta_t) \cdot (X'_t - X_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\Theta}_t)(X_t)] \cdot (X'_t - X_t)] dt \right. \\ &\quad \left. - \mathbb{E} \int_0^T [b(t, \theta'_t) - b(t, \theta_t)] \cdot Y_t + [\sigma(t, \theta'_t) - \sigma(t, \theta_t)] \cdot Z_t] dt \right\} \\ &\quad - \left\{ \mathbb{E} \int_0^T [(X'_t - X_t) \cdot \mathcal{I}_t^f + (\mathcal{I}_t^b - \mathcal{I}_t^{b,\prime}) \cdot Y_t + (\mathcal{I}_t^\sigma - \mathcal{I}_t^{\sigma,\prime}) \cdot Z_t] dt \right\} \\ &= T_0 - \gamma T_1 - T_2.\end{aligned}$$

Moreover, following (35),

$$\begin{aligned}\mathbb{E}[(X'_T - X_T) \cdot Y_T] &= \gamma \mathbb{E}[(\partial_x g(X_T, \mathbb{P}_{X_T}) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mathbb{P}_{X_T})(X_T)]) \cdot (X'_T - X_T)] \\ &\quad + \mathbb{E}[(\mathcal{I}_T^{g,\prime} - \mathcal{I}_T^g) \cdot Y_T] \\ &\leq \gamma \mathbb{E}[g(X'_T, \mathbb{P}_{X'_T}) - g(X_T, \mathbb{P}_{X_T})] + \mathbb{E}[(\mathcal{I}_T^{g,\prime} - \mathcal{I}_T^g) \cdot Y_T].\end{aligned}$$

Identifying the two expressions right above and then repeating the proof of Theorem 4.6, we obtain

$$(47) \quad \gamma J(\alpha') - \gamma J(\alpha) \geq \gamma \lambda \mathbb{E} \int_0^T |\alpha_t - \alpha'_t|^2 dt + T_0 - T_2 + \mathbb{E}[(\mathcal{I}_T^g - \mathcal{I}_T^{g,\prime}) \cdot Y_T].$$

Now, we can reverse the roles of α and α' in (47). Denoting by T'_0 and T'_2 the corresponding terms in the inequality and then making the sum of both inequalities, we deduce that:

$$2\gamma \lambda \mathbb{E} \int_0^T |\alpha_t - \alpha'_t|^2 dt + T_0 + T'_0 - (T_2 + T'_2) + \mathbb{E}[(\mathcal{I}_T^g - \mathcal{I}_T^{g,\prime}) \cdot (Y_T - Y'_T)] \leq 0.$$

The sum $T_2 + T'_2$ reads

$$\begin{aligned}&T_2 + T'_2 \\ &= \mathbb{E} \int_0^T [-(\mathcal{I}_t^f - \mathcal{I}_t^{f,\prime}) \cdot (X_t - X'_t) + (\mathcal{I}_t^b - \mathcal{I}_t^{b,\prime}) \cdot (Y_t - Y'_t) + (\mathcal{I}_t^\sigma - \mathcal{I}_t^{\sigma,\prime}) \cdot (Z_t - Z'_t)] dt.\end{aligned}$$

Similarly,

$$T_0 + T'_0 = -\mathbb{E}[(\xi - \xi') \cdot (Y_0 - Y'_0)].$$

Therefore, using Young's inequality, there exists a constant C (the value of which may increase from line to line), C being independent of γ , such that, for any $\varepsilon > 0$,

$$(48) \quad \gamma \mathbb{E} \int_0^T |\alpha_t - \alpha'_t|^2 dt \leq \varepsilon \|\Theta - \Theta'\|_{\mathbb{S}}^2 + \frac{C}{\varepsilon} (\mathbb{E}[|\xi - \xi'|^2] + \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{I}}^2).$$

Now, we observe by standard estimates for BSDEs that there exists a constant C , independent of γ , such that

$$\begin{aligned}(49) \quad &\mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t - Y'_t|^2 + \int_0^T |Z_t - Z'_t|^2 dt \right] \\ &\leq C \gamma \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X'_t|^2 + \int_0^T |\alpha_t - \alpha'_t|^2 dt \right] + C \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{I}}^2.\end{aligned}$$

Similarly,

$$(50) \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X'_t|^2 \right] \leq \mathbb{E}[|\xi - \xi'|^2] + C \gamma \mathbb{E} \int_0^T |\alpha_t - \alpha'_t|^2 dt + C \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{I}}^2.$$

From (49) and (50) and then from (48), we deduce that

$$\begin{aligned}
 (51) \quad & \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X'_t|^2 + \sup_{0 \leq t \leq T} |Y_t - Y'_t|^2 + \int_0^T |Z_t - Z'_t|^2 dt \right] \\
 & \leq C\gamma \mathbb{E} \int_0^T |\alpha_t - \alpha'_t|^2 dt + C(\mathbb{E}[|\xi - \xi'|^2] + \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{H}}^2) \\
 & \leq C\varepsilon \|\Theta - \Theta'\|_{\mathbb{S}}^2 + \frac{C}{\varepsilon} (\mathbb{E}[|\xi - \xi'|^2] + \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{H}}^2).
 \end{aligned}$$

Using the Lipschitz property of $\hat{\alpha}(t, \cdot, \cdot, \cdot, \cdot)$ and then choosing ε small enough, we complete the proof.

5.6. Decoupling Field. The notion of decoupling field, also referred to as ‘FBSDE value function’, plays a main role in the machinery of forward-backward equations, since it permits to represent the value Y_t of the backward process at time t as a function of the value X_t of the forward process at time t . When the coefficients of the forward-backward equation are random, the decoupling field is a random field. When the coefficients are deterministic, the decoupling field is a deterministic function, which solves some corresponding partial differential equation. Here is the structure of the decoupling field in the McKean-Vlasov framework:

Lemma 5.6. *For any $t \in [0, T]$ and any $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$, there exists a unique solution, denoted by $(X_s^{t,\xi}, Y_s^{t,\xi}, Z_s^{t,\xi})_{t \leq s \leq T}$, of (42) when set on $[t, T]$ with $X_t^{t,\xi} = \xi$ as initial condition.*

In this framework, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a measurable mapping $u(t, \cdot, \mu) : \mathbb{R}^d \ni x \mapsto u(t, x, \mu)$ such that

$$(52) \quad \mathbb{P}(Y_t^{t,\xi} = u(t, \xi, \mathbb{P}_\xi)) = 1.$$

Moreover, there exists a constant C , only depending on the parameters in (B1-4), such that, for any $t \in [0, T]$ and any $\xi^1, \xi^2 \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$,

$$(53) \quad \mathbb{E}[|u(t, \xi^1, \mathbb{P}_{\xi^1}) - u(t, \xi^2, \mathbb{P}_{\xi^2})|^2] \leq C\mathbb{E}[|\xi^1 - \xi^2|^2].$$

The proof is given right below. For the moment, we notice that the additional variable \mathbb{P}_ξ is for free in the above writing since we could set $v(t, \cdot) = u(t, \cdot, \mathbb{P}_\xi)$ and then have $Y_t^{t,\xi} = v(t, \xi)$. The additional variable \mathbb{P}_ξ is specified to emphasize the non-Markovian nature of the equation over the state space \mathbb{R}^d : starting from two different initial conditions, the decoupling fields might not be the same, since the law of the initial conditions might be different. Keep indeed in mind that, in the Markovian framework, the decoupling field is the same for all possible initial conditions, thus yielding the connection with partial differential equations. Here the Markov property holds, but over the enlarged space $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, thus justifying the use of the extra variable \mathbb{P}_ξ . Nevertheless, we often forget to specify the dependence upon \mathbb{P}_ξ in the sequel of the paper.

An important fact is that the representation formula (52) can be extended to the whole path:

Proposition 5.7. *Under (B1-4), for any $\xi \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$, there exists a measurable mapping $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that*

$$\mathbb{P}(\forall t \in [0, T], Y_t^{0,\xi} = v(t, X_t^{0,\xi})) = 1.$$

It satisfies $\sup_{0 \leq t \leq T} |v(t, 0)| < +\infty$. Moreover, there exists a constant C such that $v(t, \cdot)$ is C -Lipschitz continuous for any $t \in [0, T]$.

We start with

Proof of Lemma 5.6. Given $t \in [0, T]$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$, existence and uniqueness to (42) when set on $[t, T]$ with ξ as initial condition is a direct consequence of Theorem 5.1 (or, more precisely, of the proof of it since we are handling a random initial condition). Using as underlying filtration the augmented filtration \mathbb{F}^t generated by ξ and by $(W_s - W_t)_{t \leq s \leq T}$, we deduce that $Y_t^{t, \xi}$ coincides a.s. with a $\sigma(\xi)$ -measurable \mathbb{R}^d -valued random variable. In particular, there exists a measurable function $u_\xi(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\mathbb{P}(Y_t^{t, \xi} = u_\xi(t, \xi)) = 1$.

We now claim that the law of $(\xi, Y_t^{t, \xi})$ only depends upon the law of ξ . This directly follows from the version of the Yamada-Watanabe theorem for FBSDEs, see [9]: Since uniqueness holds pathwise, it also holds in law so that, given two initial conditions with the same law, the solutions also have the same laws. Therefore, given another \mathbb{R}^d -valued random vector ξ' with the same law as ξ , it holds $(\xi, u_\xi(t, \xi)) \sim (\xi', u_{\xi'}(t, \xi'))$. In particular, for any measurable function $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$, the random variables $u_\xi(t, \xi) - v(\xi)$ and $u_{\xi'}(t, \xi') - v(\xi')$ have the same law. Choosing $v = u_\xi(t, \cdot)$, we deduce that $u_{\xi'}(t, \cdot)$ and $u_\xi(t, \cdot)$ are a.e. equal under the probability measure \mathbb{P}_ξ . Put it differently, denoting by μ the law of ξ , there exists an element $u(t, \cdot, \mu) \in L^2(\mathbb{R}^d, \mu)$ such that $u_\xi(t, \cdot)$ and $u_{\xi'}(t, \cdot)$ coincide μ a.e. with $u(t, \cdot, \mu)$. Identifying $u(t, \cdot, \mu)$ with one of its version, this proves that

$$\mathbb{P}(Y_t^{t, \xi} = u(t, \xi, \mu)) = 1.$$

When $t > 0$, we notice that, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, there exists an \mathcal{F}_t -measurable random variable ξ such that $\mu = \mathbb{P}_\xi$. In such a case, the procedure we just described permits to define $u(t, \cdot, \mu)$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. The situation may be different when $t = 0$ as \mathcal{F}_0 may reduce to events of measure zero or one. In such a case, \mathcal{F}_0 can be enlarged without any loss of generality in order to support \mathbb{R}^d -valued random variables of any arbitrarily prescribed distribution.

The Lipschitz property (53) of $u(0, \cdot, \cdot)$ is a direct consequence of Lemma 5.5 with $\gamma = 1$. By a time shift, the same argument applies to $u(t, \cdot, \cdot)$. \square

We now turn to

Proof of Proposition 5.7. For simplicity, we just denote $(X_t^{0, \xi}, Y_t^{0, \xi}, Z_t^{0, \xi})_{0 \leq t \leq T}$ by $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$. The proof is then a combination of Lemmas 3.2 and 5.6. Indeed, given $t \in (0, T]$, Lemma 5.6 says that the family $(u(t, \cdot, \mu))_{\mu \in \mathcal{P}_2(\mathbb{R}^d)}$ satisfies (8) since any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ can be seen as the law of some \mathcal{F}_t -measurable random vector ζ . Therefore, for $\mu = \mathbb{P}_{X_t}$, we can find a mapping $w(t, \cdot)$ that is C -Lipschitz continuous (for the same C as in (53)) and that coincides with $u(t, \cdot, \mathbb{P}_{X_t})$ a.e. under the probability measure \mathbb{P}_{X_t} . It satisfies

$$(54) \quad \forall t \in [0, T], \quad \mathbb{P}(Y_t = w(t, X_t)) = 1,$$

since $Y_t = Y_t^{t, X_t}$. In particular,

$$(55) \quad \begin{aligned} \sup_{0 \leq t \leq T} |w(t, 0)| &\leq \sup_{0 \leq t \leq T} \mathbb{E}[|Y_t|] + \sup_{0 \leq t \leq T} \mathbb{E}[|w(t, X_t) - w(t, 0)|] \\ &\leq \sup_{0 \leq t \leq T} \mathbb{E}[|Y_t|] + C \sup_{0 \leq t \leq T} \mathbb{E}[|X_t|] < +\infty. \end{aligned}$$

For any integer $n \geq 1$, we then let

$$v^n(t, x) = \mathbf{1}_{[0, T/2^n]}(t) w\left(\frac{T}{2^n}, x\right) + \sum_{k=2}^{2^n} \mathbf{1}_{((k-1)T/2^n, kT/2^n]}(t) w\left(\frac{kT}{2^n}, x\right), \quad t \in [0, T], \quad x \in \mathbb{R}^d.$$

Denoting by $v^{n,i}$ the i th coordinate of v^n for any $i \in \{1, \dots, d\}$, we also let

$$v^i(t, x) = \limsup_{n \rightarrow +\infty} v^{n,i}(t, x), \quad t \in [0, T], \quad x \in \mathbb{R}^d,$$

and then $v(t, x) = (v^1(t, x), \dots, v^d(t, x))$. As each of the v^n is a Borel measurable function on $[0, T] \times \mathbb{R}^d$, so is v . Similarly, v satisfies (55) and, for any $t \in [0, T]$, $v(t, \cdot)$ is C -Lipschitz continuous.

Finally, we notice that, for any $t \in \mathbb{D}_n = \{kT/2^n, k \in \{1, \dots, 2^n\}\}$, with $n \in \mathbb{N} \setminus \{0\}$, and for $\ell \geq n$,

$$v^\ell(t, \cdot) = w(t, \cdot) = v(t, \cdot),$$

so that, $w(t, \cdot) = v(t, \cdot)$ for any $t \in \mathbb{D} = \cup_{n \geq 1} \mathbb{D}_n$. Therefore, \mathbb{D} being countable, we deduce from (54) that the event

$$A = \{\omega \in \Omega : \forall t \in \mathbb{D}, Y_t(\omega) = w(t, X_t(\omega)) = v(t, X_t(\omega))\}$$

has measure $\mathbb{P}(A) = 1$. On the event A , we notice that, for any $t \in (0, T]$,

$$Y_t = \lim_{n \rightarrow +\infty} Y_{t_n} = \lim_{n \rightarrow +\infty} v(t_n, X_{t_n}),$$

where $(t_n)_{n \geq 1}$ is the sequence of points in \mathbb{D} such that, for any $n \geq 1$, $t_n \in \mathbb{D}_n$ and $t_n - T/2^n < t \leq t_n$. Since $v(t_n, \cdot)$ is C -Lipschitz continuous, we deduce that, on A ,

$$Y_t = \lim_{n \rightarrow +\infty} v(t_n, X_t),$$

which is to say that the sequence $(v(t_n, X_t))_{n \geq 1}$ is convergent. Now we observe that $v(t_n, X_t)$ is also $v_n(t, X_t)$. Therefore, the limit must coincide with $v(t, X_t)$. This proves that, on the event A , $Y_t = v(t, X_t)$ for any $t \in (0, T]$. By the same argument, the same holds true at $t = 0$ (0 being handled apart for questions of notation since the definition of v_n at time 0 is rather specific). \square

6. PROPAGATION OF CHAOS AND APPROXIMATE EQUILIBRIUM

In this section, we show how the solution of the optimal control of stochastic dynamics of the McKean-Vlasov type can be used to handle N -player games when N tends to $+\infty$.

Throughout this section, assumptions (B1-4) are in force. For each integer $N \geq 1$, we consider a stochastic system whose time evolution is given by a system of N coupled stochastic differential equations of the form

$$(56) \quad dU_t^i = b(t, U_t^i, \bar{v}_t^N, \beta_t^i) dt + \sigma(t, U_t^i, \bar{v}_t^N, \beta_t^i) dW_t^i, \quad 1 \leq i \leq N; \quad \bar{v}_t^N = \frac{1}{N} \sum_{j=1}^N \delta_{U_t^j},$$

with $t \in [0, T]$ and $U_0^i = x_0$, $1 \leq i \leq N$. Here $((\beta_t^i)_{0 \leq t \leq T})_{1 \leq i \leq N}$ are N \mathbb{R}^k -valued processes that are progressively measurable with respect to the filtration generated by (W^1, \dots, W^N) and have

finite L^2 norms over $[0, T] \times \Omega$:

$$\forall i \in \{1, \dots, N\}, \quad \mathbb{E} \int_0^T |\beta_t^i|^2 dt < +\infty,$$

where, for convenience, we have fixed an infinite sequence $((W_t^i)_{0 \leq t \leq T})_{i \geq 1}$ of independent m -dimensional Brownian motions. One should think of U_t^i as the (private) state at time t of agent or player $i \in \{1, \dots, N\}$, β_t^i being the action taken at time t by player i . For each $1 \leq i \leq N$, we denote by

$$(57) \quad J^{N,i}(\beta^1, \dots, \beta^N) = \mathbb{E} \left[g(U_T^i, \bar{v}_T^N) + \int_0^T f(t, U_t^i, \bar{v}_t^N, \beta_t^i) dt \right],$$

the cost to the i th player. We then recover the same set-up as in the case of the mean field game models studied in [7]. Anyhow, the rule we apply for minimizing the cost is a bit different. The point is indeed to minimize the cost over exchangeable strategies: when the family $(\beta^i, W^i)_{1 \leq i \leq N}$ is exchangeable (with an abusive terminology, we will say that the strategy is exchangeable), the costs to all the players are the same and thus read as a common cost $J^{N,i}(\beta) = J^N(\beta)$. From a practical point of view, restricting the minimization to exchangeable strategies means that the players are intended to obey a common policy, which is not the case in the standard mean field game approach.

In this framework, one of our goal is to compute the limit

$$\lim_{N \rightarrow +\infty} \inf_{\underline{\beta}} J^N(\underline{\beta}),$$

the infimum being taken over exchangeable strategies. Another one is to identify, for each integer N , a specific set of ε -optimal strategies and the corresponding state evolutions.

6.1. Limit of the Costs and Non-Markovian Approximate Equilibriums. Recall that we denote by J the optimal cost:

$$(58) \quad J = \mathbb{E} \left[g(X_T, \mu_T) + \int_0^T f(t, X_t, \mu_t, \hat{\alpha}(t, X_t, \mu_t, Y_t, Z_t)) dt \right],$$

where $(X_t, Y_t, Z_t)_{0 \leq t \leq T}$ is the solution to (42) with $X_0 = x_0$ as initial condition, $(\mu_t)_{0 \leq t \leq T}$ denoting the flow of marginal probability measures $\mu_t = \mathbb{P}_{X_t}$, for $0 \leq t \leq T$.

For the purpose of comparison, we introduce $(\bar{X}^1, \dots, \bar{X}^N)$, each \bar{X}^i standing for the solution of the forward equation in (42) when driven by the Brownian motion W^i . Put it differently, $(\bar{X}^1, \dots, \bar{X}^N)$ solves the system (56) when the empirical distribution \bar{v}_t^N is replaced by μ_t and β_t^i is given by $\beta_t^i = \bar{\alpha}_t^i$ with

$$\bar{\alpha}_t^i = \hat{\alpha}(t, \bar{X}_t^i, \mu_t, \bar{Y}_t^i, \bar{Z}_t^i),$$

the pair (\bar{Y}^i, \bar{Z}^i) solving the backward equation in (42) when driven by W^i . Pay attention that the processes $((\bar{\Theta}_t^i = (\bar{X}_t^i, \mu_t, \bar{Y}_t^i, \bar{Z}_t^i, \bar{\alpha}_t^i))_{0 \leq t \leq T})_{1 \leq i \leq N}$ are independent.

Here is the first result:

Theorem 6.1. *Under assumptions (B1-4),*

$$\lim_{N \rightarrow +\infty} \inf_{\underline{\beta}} J^N(\underline{\beta}) = J,$$

the infimum being taken over (square integrable) strategies $\underline{\beta} = (\beta^1, \dots, \beta^N)$ such that the family $(\beta^i, W^i)_{1 \leq i \leq N}$ is exchangeable. Moreover, the non-Markovian control $\underline{\bar{\alpha}} = (\bar{\alpha}^1, \dots, \bar{\alpha}^N)$ is an approximate optimal control in the sense that

$$\lim_{N \rightarrow +\infty} J^N(\underline{\bar{\alpha}}) = J.$$

Proof. The proof consists in comparing $J^N(\underline{\beta})$ to J for a given exchangeable strategy $\underline{\beta}$. Once again, it relies on a variant of the stochastic maximum principle exposed in Section 4. With the above notation, we indeed obtain

$$J^N(\underline{\beta}) - J = \mathbb{E}[g(U_T^i, \bar{\nu}_T^N) - g(\bar{X}_T^i, \mu_T)] + \mathbb{E}\left[\int_0^T (f(s, U_s^i, \bar{\nu}_s^N, \beta_s^i) - f(s, \bar{X}_s^i, \mu_s, \bar{\alpha}_s^i)) ds\right],$$

the identity holding true for any $1 \leq i \leq N$. Therefore, we can write

$$(59) \quad J^N(\underline{\beta}) - J = T_1^i + T_2^i,$$

with

$$\begin{aligned} T_1^i &= \mathbb{E}[(U_T^i - \bar{X}_T^i) \cdot \bar{Y}_T^i] + \mathbb{E}\left[\int_0^T (f(s, U_s^i, \bar{\nu}_s^N, \beta_s^i) - f(s, \bar{X}_s^i, \mu_s, \bar{\alpha}_s^i)) ds\right], \\ T_2^i &= \mathbb{E}[g(U_T^i, \bar{\nu}_T^N) - g(\bar{X}_T^i, \mu_T)] - \mathbb{E}[(U_T^i - \bar{X}_T^i) \cdot \partial_x g(\bar{X}_T^i, \mu_T)] \\ &\quad - \mathbb{E}\tilde{\mathbb{E}}[(\tilde{U}_T^i - \tilde{\bar{X}}_T^i) \cdot \partial_\mu g(\bar{X}_T^i, \mu_T)(\tilde{\bar{X}}_T^i)] \\ &= T_{2,1}^i - T_{2,2}^i - T_{2,3}^i, \end{aligned}$$

where we used Fubini's theorem with the independent copies denoted with a tilde ' \sim '.

Analysis of T_2^i . Using the diffusive effect of independence, we claim

$$\begin{aligned} T_{2,3}^i &= \mathbb{E}\tilde{\mathbb{E}}[(\tilde{U}_T^i - \tilde{\bar{X}}_T^i) \cdot \partial_\mu g(\bar{X}_T^i, \mu_T)(\tilde{\bar{X}}_T^i)] \\ &= \frac{1}{N} \sum_{j=1}^N \tilde{\mathbb{E}}[(\tilde{U}_T^i - \tilde{\bar{X}}_T^i) \cdot \partial_\mu g(\bar{X}_T^j, \mu_T)(\tilde{\bar{X}}_T^i)] \\ &\quad + \mathcal{O}\left(\tilde{\mathbb{E}}[|\tilde{U}_T^i - \tilde{\bar{X}}_T^i|^2]^{1/2} \tilde{\mathbb{E}}\left[\left|\frac{1}{N} \sum_{j=1}^N \partial_\mu g(\bar{X}_T^j, \mu_T)(\tilde{\bar{X}}_T^i) - \mathbb{E}[\partial_\mu g(\bar{X}_T^i, \mu_T)(\tilde{\bar{X}}_T^i)]\right|^2\right]^{1/2}\right) \\ &= \frac{1}{N} \sum_{j=1}^N \mathbb{E}[(U_T^i - \bar{X}_T^i) \cdot \partial_\mu g(\bar{X}_T^j, \mu_T)(\bar{X}_T^i)] + \mathbb{E}[|U_T^i - \bar{X}_T^i|^2]^{1/2} \mathcal{O}(N^{-1/2}), \end{aligned}$$

where $\mathcal{O}(\cdot)$ stands for the Landau notation. Therefore, taking advantage of the exchangeability in order to handle the remainder, we obtain

$$\frac{1}{N} \sum_{i=1}^N T_{2,3}^i = \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N \mathbb{E}[(U_T^i - \bar{X}_T^i) \cdot \partial_\mu g(\bar{X}_T^j, \mu_T)(\bar{X}_T^i)] + \mathbb{E}[|U_T^1 - \bar{X}_T^1|^2]^{1/2} \mathcal{O}(N^{-1/2}).$$

Introducing a random variable ϑ from $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ into \mathbb{R} with uniform distribution on $\{1, \dots, N\}$ as done in the proof of Proposition 3.1, we can write

$$\frac{1}{N} \sum_{i=1}^N T_{2,3}^i = \frac{1}{N} \sum_{j=1}^N \mathbb{E} \tilde{\mathbb{E}}[(U_T^\vartheta - \bar{X}_T^\vartheta) \cdot \partial_{\mu} g(\bar{X}_T^j, \mu_T)(\bar{X}_T^\vartheta)] + \mathbb{E}[|U_T^1 - \bar{X}_T^1|^2]^{1/2} \mathcal{O}(N^{-1/2}).$$

Finally, defining the flow of empirical measures

$$\bar{\mu}_t^N = \frac{1}{N} \sum_{j=1}^N \delta_{\bar{X}_t^j}, \quad t \in [0, T],$$

and using (B3), Propositions 3.1 and 3.3 and also Remark 3.4 to estimate the distance $W_2(\bar{\mu}_T^N, \mu_T)$, the above estimate gives:

$$\frac{1}{N} \sum_{i=1}^N T_{2,3}^i = \frac{1}{N} \sum_{j=1}^N \mathbb{E} \tilde{\mathbb{E}}[(U_T^\vartheta - \bar{X}_T^\vartheta) \cdot \partial_{\mu} g(\bar{X}_T^j, \bar{\mu}_T^N)(\bar{X}_T^\vartheta)] + \mathbb{E}[|U_T^1 - \bar{X}_T^1|^2]^{1/2} \mathcal{O}(\ell_N(d)),$$

where we used the notation $\ell_N(d)$ for any function of N which could be used as an upper bound for:

$$(60) \quad \mathbb{E}[W_2^2(\bar{\mu}_T^N, \mu_T)]^{1/2} + \left(\int_0^T \mathbb{E}[W_2^2(\bar{\mu}_t^N, \mu_t)] dt \right)^{1/2} = \mathcal{O}(\ell_N(d)).$$

By Remark 3.4, the left-hand side tends to 0 as N tends to $+\infty$, since the function $[0, T] \ni t \mapsto \mathbb{E}[W_2^2(\bar{\mu}_t^N, \mu_t)]$ can be bounded independently of N . Therefore, $(\ell_N(d))_{N \geq 1}$ is always chosen as a sequence that converges to 0 as N tends to $+\infty$. When $\sup_{0 \leq t \leq T} |\bar{X}_t^1|$ has finite moment of order $d+5$, Remark 3.4 says that $\ell_N(d)$ can be chosen as $N^{-1/(d+4)}$. In any case, we will assume that $\ell_N(d) \geq N^{-1/2}$. Going back to (59),

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N T_2^i &= \frac{1}{N} \sum_{i=1}^N \left\{ \mathbb{E}[g(U_T^i, \bar{\nu}_T^N) - g(\bar{X}_T^i, \bar{\mu}_T^N)] - \mathbb{E}[(U_T^i - \bar{X}_T^i) \cdot \partial_x g(\bar{X}_T^i, \bar{\mu}_T^N)] \right. \\ &\quad \left. - \mathbb{E} \tilde{\mathbb{E}}[(U_T^\vartheta - \bar{X}_T^\vartheta) \cdot \partial_{\mu} g(\bar{X}_T^i, \bar{\mu}_T^N)(\bar{X}_T^\vartheta)] \right\} + (1 + \mathbb{E}[|U_T^1 - \bar{X}_T^1|^2]^{1/2}) \mathcal{O}(\ell_N(d)), \end{aligned}$$

where we used local Lipschitz property of g and Remark 3.4 to replace μ_T by $\bar{\mu}_T^N$.

Noting that a.s. under \mathbb{P} , the law of U_T^ϑ (resp. \bar{X}_T^ϑ) under $\tilde{\mathbb{P}}$ is the empirical distribution $\bar{\nu}_T^N$ (resp. $\bar{\mu}_T^N$), we can apply the convexity property of g , see (13), to get

$$(61) \quad \frac{1}{N} \sum_{i=1}^N T_2^i \geq -(1 + \mathbb{E}[|U_T^1 - \bar{X}_T^1|^2]^{1/2}) \mathcal{O}(\ell_N(d)).$$

Analysis of T_1^i . Using Itô's formula and Fubini's theorem, we obtain

$$\begin{aligned}
 (62) \quad T_1^i &= \mathbb{E} \left[\int_0^T (H(s, U_s^i, \bar{\nu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \beta_s^i) - H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)) ds \right] \\
 &\quad - \mathbb{E} \left[\int_0^T (U_s^i - \bar{X}_s^i) \cdot \partial_x H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i) ds \right] \\
 &\quad - \mathbb{E} \tilde{\mathbb{E}} \left[\int_0^T (\tilde{U}_s^i - \tilde{X}_s^i) \cdot \partial_\mu H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)(\tilde{X}_s^i) ds \right] \\
 &= T_{1,1}^i - T_{1,2}^i - T_{1,3}^i.
 \end{aligned}$$

Using the local Lipschitz property of the Hamiltonian and (60) and recalling that the limit process $(\bar{X}_t^i, \mu_t, \bar{Y}_t^i, \bar{Z}_t^i, \bar{\alpha}_t^i)_{0 \leq t \leq T}$ has finite \mathbb{S} -norm (see (43)), we get:

$$T_{1,1}^i = \mathbb{E} \left[\int_0^T (H(s, U_s^i, \bar{\nu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \beta_s^i) - H(s, \bar{X}_s^i, \bar{\mu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)) ds \right] + \mathcal{O}(\ell_N(d)).$$

Similarly, by exchangeability:

$$T_{1,2}^i = \mathbb{E} \left[\int_0^T (U_s^i - \bar{X}_s^i) \cdot \partial_x H(s, \bar{X}_s^i, \bar{\mu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i) ds \right] + \left(\mathbb{E} \int_0^T |U_s^i - \bar{X}_s^i|^2 ds \right)^{1/2} \mathcal{O}(\ell_N(d)).$$

Finally, using the diffusive effect of independence, we have

$$\begin{aligned}
 (63) \quad \frac{1}{N} \sum_{i=1}^N T_{1,3}^i &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \tilde{\mathbb{E}} \left[\int_0^T (U_s^i - \bar{X}_s^i) \cdot \partial_\mu H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)(\bar{X}_s^i) ds \right] \\
 &= \frac{1}{N^2} \sum_{j=1}^N \sum_{i=1}^N \mathbb{E} \left[\int_0^T (U_s^i - \bar{X}_s^i) \cdot \partial_\mu H(s, \bar{X}_s^j, \mu_s, \bar{Y}_s^j, \bar{Z}_s^j, \bar{\alpha}_s^j)(\bar{X}_s^i) ds \right] \\
 &\quad + \left(\mathbb{E} \int_0^T |U_s^1 - \bar{X}_s^1|^2 ds \right)^{1/2} \mathcal{O}(N^{-1/2}).
 \end{aligned}$$

By (B3), Propositions 3.1 and 3.3, we have

$$\begin{aligned}
 \frac{1}{N} \sum_{i=1}^N T_{1,3}^i &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \tilde{\mathbb{E}} \left[\int_0^T (U_s^\vartheta - \bar{X}_s^\vartheta) \cdot \partial_\mu H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)(\bar{X}_s^\vartheta) ds \right] \\
 &\quad + \left(\mathbb{E} \int_0^T |U_s^1 - \bar{X}_s^1|^2 ds \right)^{1/2} \mathcal{O}(N^{-1/2}) \\
 &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \tilde{\mathbb{E}} \left[\int_0^T (U_s^\vartheta - \bar{X}_s^\vartheta) \cdot \partial_\mu H(s, \bar{X}_s^i, \bar{\mu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i)(\bar{X}_s^\vartheta) ds \right] \\
 &\quad + \left(\mathbb{E} \int_0^T |U_s^1 - \bar{X}_s^1|^2 ds \right)^{1/2} \mathcal{O}(\ell_N(d)).
 \end{aligned}$$

In order to complete the proof, we evaluate the missing term in the Taylor expansion of T_1^i in (62), namely

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[\int_0^T (\beta_s^i - \bar{\alpha}_s^i) \cdot \partial_\alpha H(s, \bar{X}_s^i, \bar{\mu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i) ds \right],$$

in order to benefit from the convexity of H . We use Remark 3.4 once more:

$$\begin{aligned} & \mathbb{E} \left[\int_0^T (\beta_s^i - \bar{\alpha}_s^i) \cdot \partial_\alpha H(s, \bar{X}_s^i, \bar{\mu}_s^N, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i) ds \right] \\ (64) \quad &= \mathbb{E} \left[\int_0^T (\beta_s^i - \bar{\alpha}_s^i) \cdot \partial_\alpha H(s, \bar{X}_s^i, \mu_s, \bar{Y}_s^i, \bar{Z}_s^i, \bar{\alpha}_s^i) ds \right] + \left(\mathbb{E} \int_0^T |\beta_s^i - \bar{\alpha}_s^i|^2 ds \right)^{1/2} \mathcal{O}(\ell_N(d)) \\ &= \left(\mathbb{E} \int_0^T |\beta_s^i - \bar{\alpha}_s^i|^2 ds \right)^{1/2} \mathcal{O}(\ell_N(d)), \end{aligned}$$

since $\bar{\alpha}$ is an optimizer for H . Using the convexity of H and taking advantage of the exchangeability, we finally deduce from (62), (63) and (64) that there exists a constant $c > 0$ such that

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N T_1^i &\geq c \mathbb{E} \int_0^T |\beta_s^1 - \bar{\alpha}_s^1|^2 ds \\ &\quad - \mathcal{O}(\ell_N(d)) \left(1 + \sup_{0 \leq t \leq T} \mathbb{E}[|U_t^1 - \bar{X}_t^1|^2] + \mathbb{E} \int_0^T |\beta_s^1 - \bar{\alpha}_s^1|^2 ds \right)^{1/2}. \end{aligned}$$

By (61) and (59), we deduce that

$$\begin{aligned} J^N(\underline{\beta}) &\geq J + c \mathbb{E} \int_0^T |\beta_s^1 - \bar{\alpha}_s^1|^2 ds \\ &\quad - \mathcal{O}(\ell_N(d)) \left(1 + \sup_{0 \leq t \leq T} \mathbb{E}[|U_t^1 - \bar{X}_t^1|^2] + \mathbb{E} \int_0^T |\beta_s^1 - \bar{\alpha}_s^1|^2 ds \right)^{1/2}. \end{aligned}$$

From the inequality

$$\sup_{0 \leq t \leq T} \mathbb{E}[|U_t^1 - \bar{X}_t^1|^2] \leq C \mathbb{E} \int_0^T |\beta_s^1 - \bar{\alpha}_s^1|^2 ds,$$

which holds for some constant C independent of N , we deduce that

$$(65) \quad J^N(\underline{\beta}) \geq J - C \ell_N(d),$$

for a possibly new value of C , which proves that

$$\liminf_{N \rightarrow +\infty} \inf_{\underline{\beta}} J^N(\underline{\beta}) \geq J.$$

In order to prove Theorem 6.1, it thus remains to find a sequence of controls $(\underline{\beta}^N)_{N \geq 1}$ such that

$$\limsup_{N \rightarrow +\infty} J^N(\underline{\beta}^N) \leq J.$$

Precisely, we show right below that

$$\limsup_{N \rightarrow +\infty} J^N(\bar{\alpha}) \leq J,$$

thus proving that $\bar{\alpha} = (\bar{\alpha}^1, \dots, \bar{\alpha}^N)$ is an approximate equilibrium, but of non-Markovian type. Denoting by (X^1, \dots, X^N) the solution of (56) with $\beta_t^i = \bar{\alpha}_t^i$, classical estimates from the theory of propagation of chaos imply (see e.g. [16] or [10]) that

$$\sup_{0 \leq t \leq T} \mathbb{E}[|X_t^i - \bar{X}_t^i|^2] = \sup_{0 \leq t \leq T} \mathbb{E}[|X_t^1 - \bar{X}_t^1|^2] = \mathcal{O}(N^{-1}).$$

It is then plain to deduce that

$$\limsup_{N \rightarrow +\infty} J^N(\bar{\alpha}) \leq J.$$

This completes the proof. \square

6.2. Approximate Equilibriums with Distributed Closed Loop Controls. When σ doesn't depend upon α , we are able to provide an approximate equilibrium using only distributed controls in closed loop form. This is of real interest from the practical point of view. Indeed, in a such case, the optimizer $\hat{\alpha}$ of the Hamiltonian, as defined in (39), doesn't depend on z . It thus reads as $\hat{\alpha}(t, x, \mu, y)$. By Proposition 5.7, this says that the optimal control $(\alpha_t)_{0 \leq t \leq T}$ in Theorem 5.1 has the *feedback* form:

$$(66) \quad \alpha_t = \hat{\alpha}(t, X_t, \mu_t, v(t, X_t)), \quad t \in [0, T].$$

The reader might object that, also in the case when σ depends upon α , the process Z_t at time t is also expected to read as a function of X_t , since such a representation is known to hold in the classical *decoupled* forward-backward setting. Even if we feel that it is indeed possible to prove such a representation in our more general setting, we must address the following points: (i) From a practical point of view, (66) is meaningful if the feedback function is Lipschitz-continuous, as the Lipschitz property ensures that the stochastic differential equation obtained by plugging (66) into the forward equation in (42) is solvable; (ii) In the current framework, the function v is known to be Lipschitz continuous by Proposition 5.7, but proving the same result for the representation of Z_t in terms of X_t seems to be really challenging (by the way, it is already challenging in the standard case, i.e. without any McKean-Vlasov interaction); (iii) We finally mention that, in any case, the relationship between Z_t and X_t , if exists, must be rather intricate as Z_t is expected to solve the equation $Z_t = \partial_x v(t, X_t) \sigma(t, X_t, \mathbb{P}_{X_t}, \hat{\alpha}(t, X_t, \mathbb{P}_{X_t}, Y_t, Z_t))$, which can be formally derived by identifying martingale integrands when expanding $Y_t = v(t, X_t)$ by a formal application of Itô's formula. This equation has been investigated in [18] in the standard case, but we feel more convenient not to repeat this analysis in the current setting in order to keep things at a reasonable level of complexity.

Now, for each integer N , we can consider the solution $(X_t^1, \dots, X_t^N)_{0 \leq t \leq T}$ of the system of N stochastic differential equations

$$(67) \quad dX_t^i = b(t, X_t^i, \mu_t^N, \hat{\alpha}(t, X_t^i, \mu_t, v(t, X_t^i)))dt + \sigma(t, X_t^i, \mu_t^N)dW_t^i, \quad \mu_t^N = \frac{1}{N} \sum_{j=1}^N \delta_{X_t^j},$$

with $t \in [0, T]$ and $X_0^i = x_0$. The system (67) is well posed since v satisfies Proposition 5.7 and the minimizer $\hat{\alpha}(t, x, \mu, y)$ is Lipschitz continuous and at most of linear growth in the variables x, μ and y , uniformly in $t \in [0, T]$. The processes $(X^i)_{1 \leq i \leq N}$ give the dynamics of the private states of the N players in the stochastic differential game of interest when the players use the strategies

$$(68) \quad \alpha_t^{N,i} = \hat{\alpha}(t, X_t^i, \mu_t, v(t, X_t^i)), \quad 0 \leq t \leq T, \quad i \in \{1, \dots, N\}.$$

These strategies are in closed loop form. They are even *distributed* since, at each time $t \in [0, T]$, a player only needs to know the state of his own private state in order to compute the value of the action to take at that time. By the linear growth of v and of the minimizer $\hat{\alpha}$, it holds, for any $p \geq 2$,

$$(69) \quad \sup_{N \geq 1} \max_{1 \leq i \leq N} [\mathbb{E}[\sup_{0 \leq t \leq T} |X_t^i|^p]] < +\infty,$$

the expectation inside the brackets being actually independent of i since the strategy is obviously exchangeable.

We then have the approximate equilibrium property:

Theorem 6.2. *In addition to assumptions (B1-4), assume that σ doesn't depend upon α . Then,*

$$J^N(\underline{\beta}) \geq J^N(\underline{\alpha}^N) - \mathcal{O}(N^{-1/(d+4)}),$$

for any $\underline{\beta} = (\beta^1, \dots, \beta^N)$ such that $(\beta^i, W^i)_{1 \leq i \leq N}$ is exchangeable, where α is defined in (68).

Proof. We use the same notations as in the proof of Theorem 6.1.

Since $\bar{\alpha}_t^1$ now reads as $\hat{\alpha}(t, \bar{X}_t^1, \mu_t, v(t, \bar{X}_t^1))$ for $0 \leq t \leq T$, we first notice, by the growth property of v , that $\mathbb{E}[\sup_{0 \leq t \leq T} |\bar{X}_t^1|^p] < +\infty$ for any $p \geq 1$. As mentioned in (11) in Remark 3.4, this says that $\ell_N(d)$ in the lower bound

$$J^N(\underline{\beta}) \geq J - C\ell_N(d),$$

see (65), can be chosen as $N^{-1/(d+4)}$.

Moreover, since $v(t, \cdot)$ is Lipschitz continuous, using once again classical estimates from the theory of propagation of chaos (see e.g. [16] or [10]), we also have

$$\sup_{0 \leq t \leq T} \mathbb{E}[|X_t^i - \bar{X}_t^i|^2] = \sup_{0 \leq t \leq T} \mathbb{E}[|X_t^1 - \bar{X}_t^1|^2] = \mathcal{O}(N^{-1}),$$

so that

$$\sup_{0 \leq t \leq T} \mathbb{E}[|\alpha_t^{N,i} - \bar{\alpha}_t^i|^2] = \sup_{0 \leq t \leq T} \mathbb{E}[|\alpha_t^{N,1} - \bar{\alpha}_t^1|^2] = \mathcal{O}(N^{-1}),$$

for any $1 \leq i \leq N$. It is then plain to deduce that

$$J^N(\alpha^N) \leq J + C\ell_N(d).$$

This completes the proof. \square

REFERENCES

- [1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the Wasserstein space of probability measures*. Birkhäuser, 2004.
- [2] D. Andersson and B. Djehiche. A maximum principle for sdes of mean-field type. *Applied Mathematics & Optimization*, 63 (3):341–356, 2010.
- [3] A. Bensoussan, K.C.J. Sung, S.C.P. Yam, and S.P. Yung. Linear quadratic mean field games. Technical report, 2011.
- [4] R. Buckdahn, B. Djehiche, J. Li, and S. Peng. Mean-field backward stochastic differential equations: A limit approach. *The Annals of Probability*, 37:1524–1565, 2009.
- [5] P. Cardaliaguet. Notes on mean field games. Technical report, 2010.
- [6] R. Carmona and F. Delarue. Mean field forward-backward stochastic differential equations. Technical report.
- [7] R. Carmona and F. Delarue. Probabilistic analysis of mean field games. Technical report, 2012.
- [8] R. Carmona, F. Delarue, and A. Lachapelle. Control of McKean-Vlasov versus Mean Field Games. *Mathematical Financial Economics*, page to appear, 2012.

- [9] F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Processes and Applications*, 99:209–286, 2002.
- [10] B. Jourdain, S. Meleard, and W. Woyczynski. Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA, Latin American Journal of Probability*, 4:1–29, 2008.
- [11] H.P. McKean. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Science*, 56:1907–1911, 1966.
- [12] H.P. McKean. Propagation of chaos for a class of nonlinear parabolic equations. *Lecture Series in Differential Equations*, 7:41–57, 1967.
- [13] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Letters*, pages 55–61, 1990.
- [14] S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37:825–843, 1999.
- [15] S.T. Rachev and L. Ruschendorf. *Mass Transportation Problems I: Theory*. Springer Verlag, 1998.
- [16] A.S. Sznitman. Topics in propagation of chaos. In *D. L. Burkholder et al. , Ecole de Probabilités de Saint Flour, XIX-1989*, volume 1464 of *Lecture Notes in Mathematics*, pages 165–251, 1989.
- [17] C. Villani. *Optimal Transport, Old and New*. Springer Verlag, 2009.
- [18] Z. Wu and Z. Yu. Probabilistic interpretation for systems of parabolic partial differential equations combined with algebra equations. Technical report, 2010.

ORFE, BENDHEIM CENTER FOR FINANCE, PRINCETON UNIVERSITY, PRINCETON, NJ 08544, USA.

E-mail address: rcarmona@princeton.edu

LABORATOIRE JEAN-ALEXANDRE DIEUDONNÉ, UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS, PARC VALROSE, 06108 CEDEX 02, NICE, FRANCE

E-mail address: Francois.Delarue@unice.fr